UNIVERSITY OF MASSACHUSETTS
Department of Public Health
Program in Biostatistics and Epidemiology

PubHlth 640W - Intermediate Biostatistics

Section 01: Off-Campus (Worcester) – 58546
Section 02: Online - 58975

Spring 2014

http://people.umass.edu/~biep640w

Instructor:
Carol Bigelow, PhD
School of Public Health
402 Arnold House
University of Massachusetts
Amherst, MA 01003
tel: 413/545-1319
fax: 413/545-1645
email: cbigelow@schoolph.umass.edu

Required Text:
Rosner, Bernard
Fundamentals of Biostatistics, *Seventh Edition*
Brooks/Cole Cengage Learning
2011
ISBN-13: 978-0-538-73349-6
ISBN-10: 0-538-73349-7
Note- Feel welcome to purchase an earlier version, as this may save you money.

Other Text Resources (NOT required)

(1) *For those interested in Stata – this is the PubHlth 691F text*
Acock AC
A Gentle Introduction to Stata, *Third Edition*
Stata Press
2010
ISBN-13: 978-1-59718-075-7
ISBN-10: 1-59718-075-0

(2) *More Advanced*
Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE
Regression Methods in Biostatistics – Linear, Logistic, Survival and Repeated Measures Models
Springer
2005
ISBN 0-387-20275-7

Statistical Software:

I am currently using Stata version 13 in my teaching and will provide illustrations of its use in this class.

How to Obtain Stata (OPTIONAL):

Stata Corp. offers student discounts on the purchase of Stata through what is called GradPlan. The cost varies, depending on the size of Stata you want (maximum number of variables, lease versus perpetual license).

Step 1: Compare features of the various options, at http://www.stata.com/products/which-stata-is-right-for-me/.

Compare features

Package	Max. no. of variables	Max. no. of right-hand variables	Max. no. of observations	64-bit version available?	Fastest: designed for parallel processing?	Platforms
Stata/MP	32,767	10,998	unlimited*	Yes	Yes	Windows, Mac, or Unix
Stata/SE	32,767	10,998	unlimited*	Yes	No	Windows, Mac, or Unix
Stata/IC	2,047	798	unlimited*	Yes	No	Windows, Mac, or Unix
Small Stata	99	98	1,200	Yes	No	Windows, Mac, or Unix

*The maximum number of observations is limited only by the amount of available RAM on your system.

Requirements

Package	Memory	Disk space
Stata/MP	512 MB	500 MB
Stata/SE	512 MB	500 MB
Stata/IC	512 MB	500 MB
Small Stata	512 MB	500 MB

Comparison:

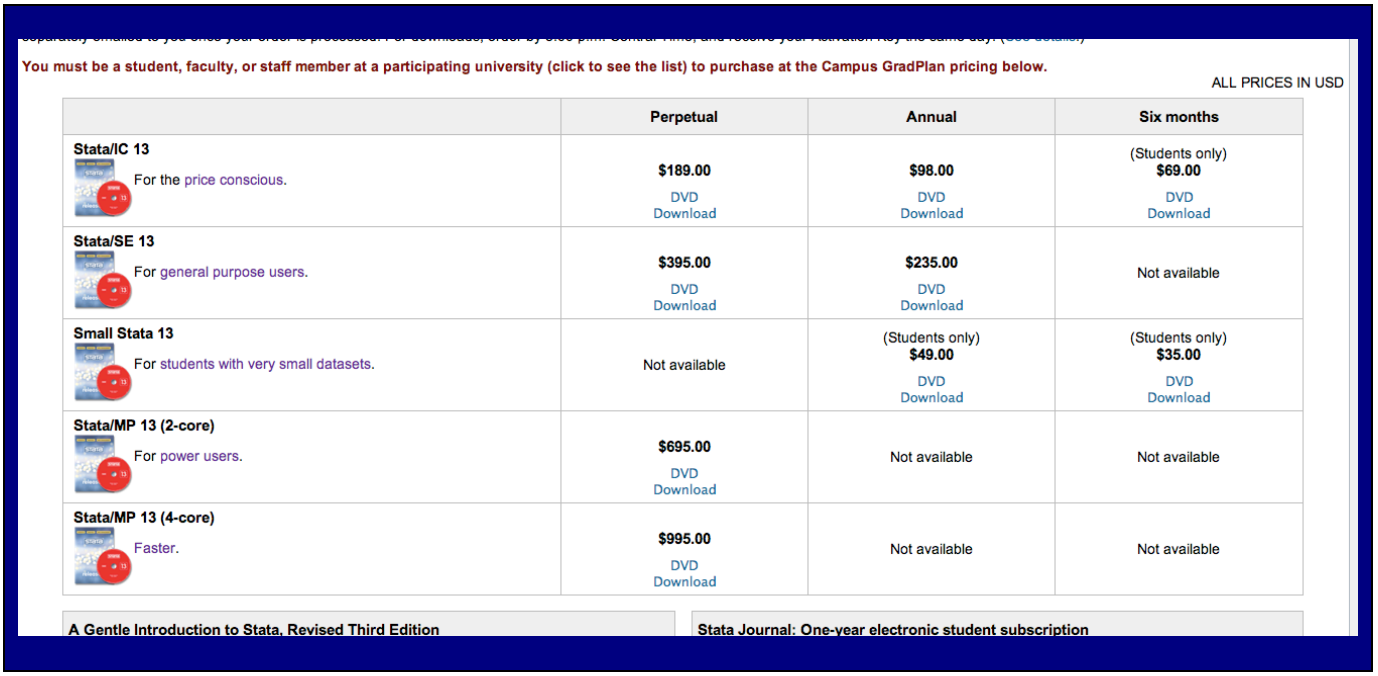
Stata/MP - Not recommended: it's expensive and far more than you will ever need

Stata/SE - Good choice if you anticipate working with very large data sets.

Stata/IC - My first choice for class work.

Small Stata - This is okay and has the advantage of being the least expensive.

Step 2: Compare prices at http://www.stata.com/order/new/edu/gradplans/.



You must be a student, faculty, or staff member at a participating university (click to see the list) to purchase at the Campus GradPlan pricing below.

ALL PRICES IN USD

	Perpetual	Annual	Six months
Stata/IC 13 For the price conscious.	**$189.00** DVD Download	**$98.00** DVD Download	(Students only) **$69.00** DVD Download
Stata/SE 13 For general purpose users.	**$395.00** DVD Download	**$235.00** DVD Download	Not available
Small Stata 13 For students with very small datasets.	Not available	(Students only) **$49.00** DVD Download	(Students only) **$35.00** DVD Download
Stata/MP 13 (2-core) For power users.	**$695.00** DVD Download	Not available	Not available
Stata/MP 13 (4-core) Faster.	**$995.00** DVD Download	Not available	Not available

A Gentle Introduction to Stata, Revised Third Edition | **Stata Journal: One-year electronic student subscription**

Step 3: You may purchase Stata online at a student discount at any time between **January 13, 2014 and May 31, 2014**.

(1) Go to the Stata Corp. GradPlan at
http://www.stata.com/order/new/edu/gradplans/course-information
(2) Have your UMass student id ready (***REQUIRED***)
(3) At the GRADPLAN ID prompt, specify: the ID (***I WILL PROVIDE THIS ELSEWHERE***)
(4) **IMPORTANT!!** BE SURE to choose the option for Stata to be shipped to your home directly.

Course Description

This course is the second of a two semester sequence (PubHlth 540 and PubHlth 640) of introductory biostatistics. The overall objective is the development of basic statistical literacy and skills in the analysis of biological and health data. Use of the computer (especially Stata) and the analysis of data sets are included. The following topics are discussed: simple linear regression, multivariable regression, analysis of proportions and rates, logistic regression, survival analysis, analysis of variance, repeated measurements analysis, and nonparametric analyses.

Course Objectives and Outcome Competencies

Course Objectives: By the end of this course, you should be able to perform, interpret, and report the findings of selected simple statistical analyses: description, hypothesis testing, simple linear regression, some multivariable regression analyses, some analysis of proportions and rates, and some analyses of variance. Time permitting, you may also be able to carry out some simple survival analyses, repeated measurements analyses and nonparametric analyses.

Outcome Competencies:

The specific outcome competencies include, but are not limited to, the following:

1. Explain the conceptual framework of selected, basic methods, of biostatistical analysis – This is "statistical literacy". In presenting each topic, I will emphasize the underlying principles, rationale, and relevance. For example, you will learn that multivariable models are likely to be wrong but that, nevertheless, a good fitted model can yield important insights into the nature and strength of associations that might exist, the latter being potentially useful in the advancement of public health.

2. Develop a conceptual framework that integrates techniques and methods in biostatistics – Two conceptual frameworks are utilized in this course. The first perspective is that the theory of biostatistics (and epidemiology, too) are about the scientific method (and the goal of causal inference). The second conceptual framework pertains to precision of estimation and statistical hypothesis testing and the notion of "signal" and "noise".

3. Integrating analysis strategies in biostatistics with principles and issues in epidemiology – Applications of biostatistics are often grounded in the tools of observational epidemiology. The presentation of the topics in this course will highlight their relevance to confounding, effect modification, discovery of intermediary pathways, and reduction of bias.

4. Apply biostatistical methods to the design of studies in public health – As the course progresses, increasingly, we will integrate basic principles of statistical literacy with those of epidemiological research to gain practice in developing data analysis plans. These vary depending on the data type and the questions of interest.

5. Use computers to appropriately store, manage, manipulate and process data for a research study using modern software – This course includes an introduction to the use of Stata for these purposes.

6. Apply descriptive techniques commonly used to summarize public health data – I will extend the introduction to descriptive techniques provided in PubHlth 540 (*Introductory Biostatistics*) to the production of "presentation" quality data summaries. I will emphasize the importance of graphical summaries and the use of Stata for data description.

7. Describe the basic concepts of probability, random variation and selected, commonly used, probability distributions – This will also be an extension of the ideas introduced in PubHlth 540 (*Introductory Biostatistics*). You will learn additional concepts of sampling distributions and additional applications of the central limit theorem. Specifically, you will learn how these ideas are the foundation of modeling, estimation, and hypothesis testing.

8. Select and perform the appropriate descriptive and inferential statistical methods in selected basic study design settings. – I will provide data sets for you to explore. I encourage you to take advantage of this opportunity to try your hand at developing your own analysis plan, doing the programming necessary for analysis, interpreting your results (especially with respect to the analysis goals and associated issues of confounding, bias, effect modification, and precision) and generating a report of your findings.

9. Interpret results and critically evaluate basic statistical aspects of public health research and practice reported in the literature – You will gain practice in being a statistically literate consumer of published examples of data analyses. You will also be encouraged to select a published article from your own particular area of interest (this might be from work or your thesis work) and writing a brief report on the nature and appropriateness of the statistics used.

10. Effective communication – The utility of biostatistics work rests, ultimately, in its effective communication. In the weekly practice assignments and in the exams, especially, you will gain practice in the communication of biostatics work to the lay reader. Specifically, you gain competencies in the following descriptions: analysis question, rationale, method used, statistical findings, and subject matter relevance.

Office Hours:
4:00-5:00 Mondays, in the hospital cafeteria, or, by appointment.

This course has 9 units
1. Review of PubHlth 540, *Introductory Biostatistics*
2. Regression and Correlation
3. Discrete Distributions
4. Categorical Data Analysis
5. Logistic Regression
6. (*Optional*) Survival Analysis
7. Analysis of Variance
8. Repeated Measures Analysis
9.(*Optional*) Nonparametrics

Grading Policy:
Your course grade will be based on completion of ***10 out of 12*** practice problems and ***three*** "take home" open book examinations, as follows.

	Percent of Course Grade
Practice Problems (10 out of 12 sets)	**25%**
Examination I (*required*)	**25%**
Examination II (*required*)	**25%**
Examination III (*required*)	**25%**

Policy on Due Dates

Sometimes things come up and it is not possible to meet a class deadline. To accommodate this, I will accept late submissions up to one week. Please be aware, however, that in consideration of your classmates, a late submission carries a 20 point penalty. Thus, if you know you cannot make a due date, your best bet is to use the full week grace time!

	Credit Policy
On Time	**Full Credit for points Scored**
1-7 Days Late	**Points Scored – 20 points**
8+ Days Late	**0 points (no credit)**

Policy on Dates of Postings

I am sorry but I do not post lecture notes or exams ahead of schedule.

Letter Grade Determination:

A	95 and over
A-	90 - 94
B+	87 - 89
B	83 – 86
B-	80 - 82
C+	77 – 79
C	70 – 76
F	Below 70

Important Dates to Remember

First Day of Spring Semester Classes:	**Tuesday January 21, 2014**
Last Day to Drop with no Record	**Monday February 3, 2014**
Holiday, President's Day – **NO CLASS**	**Monday February 17, 2014**
Monday Class Schedule will be followed	**Wednesday February 19, 2014**
Last Day to Drop with "DR"	**Monday March 3, 2014**
Spring Break Recess – **NO CLASS**	**Monday March 17, 2014**
Holiday, Patriot's Day – **NO CLASS**	**Monday April 21, 2014**
Monday Class Schedule will be followed	**Wednesday April 23, 2014**
Last Day of Spring Semester Classes:	**Wednesday April 30, 2014**
Take Home Final Exam Due	**Thursday May 8, 2014**

Schedule of PubHlth 640 Lectures and Laboratory Sessions

Meeting	Date	Lecture or Laboratory Session
1	Mon Jan 27, 2014	1 – Review of PubHlth 540, especially Simple Linear Regression
2	Mon Feb 3, 2014	2 – Regression and Correlation (Part I)
3	Mon Feb 10, 2014	2 – Regression and Correlation (Part II)
-	**Mon Feb 17, 2014**	**HOLIDAY – President's Day**
4	***Wed*** Feb 19, 2014	***STATA Lab Session*** – **Bring your laptop to class (Section 01)**
5	Mon Feb 24, 2014	3 – Discrete Distributions
6	Mon Mar 3, 2014	4 – Categorical Data Analysis (Part I)
7	Mon Mar 10, 2014	4 – Categorical Data Analysis (Part II)
-	**Mon Mar 17, 2014**	**SPRING HOLIDAY**
8	Mon Mar 24, 2014	5 – Logistic Regression (Part I)
9	Mon Mar 31, 2014	5 – Logistic Regression (Part II)
10	Mon Apr 7, 2014	6 – Introduction to Survival Analysis
11	Mon Apr 14, 2014	7 – Analysis of Variance
-	**Mon Apr 21, 2014**	**HOLIDAY – Patriot's Day**
12	***Wed*** Apr 23, 2014	***STATA Lab Session*** – **Bring your laptop to class (Section 01)**.
13	Mon Apr 28, 2014	8 – Introduction to Repeated Measures

Examination Schedule

Exam	Posting	Due	Topics Covered
1	Mon Feb 24, 2014	Mon Mar 10, 2014	1 – Review of PubHlth 540 2 – Regression and Correlation
2	Mon Mar 31, 2014	Mon Apr 14, 2014	3 – Discrete Distributions 4 – Categorical Data Analysis 5 – Logistic Regression
3	Mon April 28, 2014	Thu May 8, 2014	7 – Analysis of Variance

Note - There will be no examinations of units 6 (*Survival Analysis*), 8 (*Repeated Measurements*), or 9 (*Nonparametrics*)

Make-up and Rescheduling Policies

- I cannot promise to be able to provide all lecture notes and overheads **ahead of schedule;** sorry.
- If you miss a class, you can obtain the lecture notes from the course website http://people.umass.edu/~biep640w
- As a policy, unless there are extenuating circumstances, Linda Hollis will **not** mail out missing lecture notes and overheads.

ADA Accommodation Policy

Any student who, because of a disability, may require special arrangements in order to meet course requirements should contact me as soon as possible to make necessary arrangements.

Carol Bigelow, PhD
tel: 413/545-1319
fax: 413/545-1645
email: cbigelow@schoolph.umass.edu

Policy on Academic Dishonesty

All students are expected to adhere to guidelines of University of Massachusetts regarding academic honesty. A copy of these guidelines is available online at

www.umass.edu/dean_students/code_conduct/acad_honest.htm

The University of Massachusetts/Amherst Senate Document 89-026 defines academic dishonesty as including but not limited to:

a) Cheating – intentional deceit, trickery, or breach of confidence, used to gain some unfair or dishonest advantage in one's academic work.

b) Fabrication – intentional falsification or invention of any information or citation in any academic exercise.

c) Facilitating dishonesty – knowingly helping or attempting to help someone else commit an act of academic dishonesty.

d) Plagiarism – knowingly representing the words or ideas of another as one's own work in any academic exercise.

e) Submitting in whole or in part, without citation, prewritten term papers of another or the research of another (including but not limited to such materials sold or distributed commercially).