Epi Bio 302
Introduction to Biostatistics
Summer 2012

Instructor: Juned Siddique, Assistant Professor, 680 N. Lake Shore Dr., Suite 1400, (312) 908-9241, (siddique@northwestern.edu). Office hours: Tues. 3:00pm–4:30pm or by appointment. I will also be available directly after class to answer questions. E-mail is an effective way to reach me with requests for meetings, questions about the material, etc.

Course assistant (CA): Mr. Peter de Chavez (p-chavez@northwestern.edu). You are welcome to e-mail him with questions.

Course website: We will use the online Course Management System, Blackboard, https://courses.northwestern.edu/webapps/login/ to transmit and share all course materials including the syllabus, problem sets, data sets, and email correspondence.

Times and Place:
Lectures: Tues-Wed-Thurs 1:00–2:30pm, McGaw 2-322 (on 7/23, Class will be held in McGaw - Williams Auditorium)
Computer lab session: Thurs 3:30–5:00pm, Galter Library Learning Resources Center (LRC)
CA office hours: Tues 11am-noon and by appointment. Note: If you have a laptop that you do your work on, please feel free to bring it to meetings.

Prerequisites: College algebra and ability to use a personal computer.

Description: This course will provide an introduction to the basic concepts of statistics as applied to the biomedical and public health sciences. Emphasis is on the use and interpretation of statistical tools for data analysis. Topics include (i) descriptive statistics; (ii) probability and sampling; (iii) the methods of statistical inference; and (iv) an introduction to linear and logistic regression.

Texts: (University Bookstore)


References:
Computer Software:

Stata (Intercooled), version 11. Available in Galter library and the MSEB computer lab. Can also be purchased at: [http://www.stata.com/order/new/edu/gradplans/gp-direct.html](http://www.stata.com/order/new/edu/gradplans/gp-direct.html) A 6-month Stata IC license is $65. Please see me if you do not have access to Stata.


Homework: There will be 6 homework assignments, generally due one week after they are assigned. The final homework will cover analysis techniques from the entire course, including logistic regression. You are encouraged to work together on and discuss the homework, if doing so will help you learn the material. However, each person should carry out all analyses and write up the homework his/her self. I will try to indicate after each lecture which problems can be done at that time. If I forget to do this, please do not hesitate to remind me.

Midterm: Wednesday, 25 July. The midterm exam will be closed-book and will cover material through “Relationships among two categorical variables” (July 19).

Final: Wednesday, 22 August. The final will be closed book and will concentrate on material in the second half of the course, but not including the material on linear and logistic regression.

Grading: Homework: 60%, Midterm: 20%, Final: 20%

Objectives: The student will learn to do the following:

- Identify and give examples of nominal, ordinal, interval-scale and ratio-scale variables
- By hand and using statistical software, present and interpret empirical distribution functions for nominal, ordinal or continuous data.
- Using statistical software, graphically present the joint empirical distribution of two variables, perhaps of different types
- Using statistical software, compute and interpret measures of association of two variables, perhaps of different types
- Using the lexicon of probability, define event, probability, conditional probability, independence, mutual exclusivity and random variable
- In a biomedical or epidemiologic research study, identify the population and the sample
- Correctly interpret tests of simple hypotheses and confidence intervals
- Carry out simple power analyses and sample size calculations for one- and two-sample study designs
- Perform basic tests on association measures arising from cross-classified nominal variables
- Perform basic analyses on outcomes expressed as proportions
- Estimate and interpret the parameters in a linear regression model using statistical software
- Estimate and interpret the parameters in a logistic regression model using statistical software
- Interpret linear or logistic regression models presented in biomedical or epidemiologic research reports from scientific journals
Course Evaluations: The Programs in Public Health administer web-based course evaluations to students for each course near the end of the quarter. Your completion of both the unit (course) and faculty evaluation components is required; failure to complete either of the evaluations will result in an incomplete grade until the evaluations are submitted. You will be sent the web link and instructions via email later in the quarter. You will have about two weeks time to complete the evaluations before grades are submitted.

Academic Integrity: Every Northwestern faculty member and student belongs to a community of scholars where academic integrity is a fundamental commitment. The Program in Public Health abides by the standards of academic conduct, procedures, and sanctions as set forth by The Graduate School at Northwestern University. Students and faculty are responsible for knowledge of the information provided by The Graduate School on their Web page at: [http://www.tgs.northwestern.edu/academics/academic-services/integrity/index.html](http://www.tgs.northwestern.edu/academics/academic-services/integrity/index.html)

Academic misconduct includes, but is not limited to

1. Receiving or giving unauthorized aid on examinations or homework
2. Plagiarism
3. Fabrication
4. Falsification or manipulation of academic records
5. Aiding or abetting any of the above

The PPH follows The Graduate School’s procedure for evaluating alleged academic misconduct, as outlined on the TGS website: [http://www.tgs.northwestern.edu/academics/academic-services/integrity/dishonesty/index.html](http://www.tgs.northwestern.edu/academics/academic-services/integrity/dishonesty/index.html)

Faculty reserve the right to use the "Safe Assignment: Plagiarism Detection Tool" that is part of the Course Management System to evaluate student assignments. Information about this tool can be found at: [http://www.it.northwestern.edu/education/course-management/support/assessments/safeassignment.html](http://www.it.northwestern.edu/education/course-management/support/assessments/safeassignment.html)
<table>
<thead>
<tr>
<th>Date(s)</th>
<th>Topic</th>
<th>Reading Assignments</th>
<th>Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 3</td>
<td>Introduction</td>
<td>Ch. 1: 1.1, 1.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probability: basic concepts and definitions</td>
<td>Ch. 6: 6.1–6.4, 6.6</td>
<td></td>
</tr>
<tr>
<td>July 4</td>
<td>No class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>July 5</td>
<td>Probability (cont.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Random variables and probability distributions</td>
<td>Ch. 7: 7.1</td>
<td></td>
</tr>
<tr>
<td>July 10</td>
<td>Binomial (Bernoulli) distribution</td>
<td>Ch. 7: 7.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poisson and Normal (Gaussian) distributions</td>
<td>Ch. 7: 7.3, 7.4</td>
<td></td>
</tr>
<tr>
<td>July 11</td>
<td>Poisson and Normal (Gaussian) distributions (cont.)</td>
<td>Ch. 7: 7.3, 7.4</td>
<td></td>
</tr>
<tr>
<td>July 12</td>
<td>Data types, exploring and summarizing data</td>
<td>Ch. 2: 2.1</td>
<td>HW 1</td>
</tr>
<tr>
<td></td>
<td>Descriptive statistics for a single variable</td>
<td>Ch. 2: 2.2, 2.3; Ch. 3: 3.1–3.4</td>
<td></td>
</tr>
<tr>
<td>July 17</td>
<td>Descriptive statistics for a single variable (cont.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Relationships among two continuous variables</td>
<td>Ch. 17: 17.1, 17.2</td>
<td></td>
</tr>
<tr>
<td>July 18</td>
<td>Relationships among two continuous variables (cont.)</td>
<td>Ch. 17: 17.1, 17.2</td>
<td>HW 2</td>
</tr>
<tr>
<td></td>
<td>Relationships among two categorical variables</td>
<td>Ch. 6: 6.5; Ch. 15: 15.3</td>
<td></td>
</tr>
<tr>
<td>July 19</td>
<td>Relationships among two categorical variables (cont.)</td>
<td>Ch. 6: 6.5; Ch. 15: 15.3</td>
<td></td>
</tr>
<tr>
<td>July 24</td>
<td>Review and “Q &amp; A” Session</td>
<td></td>
<td>(cont.)</td>
</tr>
<tr>
<td>Date(s)</td>
<td>Topic</td>
<td>Reading Assignments</td>
<td>Due</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------</td>
<td>---------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>July 25</td>
<td>MID-TERM EXAM: in class</td>
<td></td>
<td>HW 3</td>
</tr>
<tr>
<td>July 26</td>
<td>Populations, studies, samples (data) and sampling distributions</td>
<td>Ch. 8: 8.1–8.4</td>
<td></td>
</tr>
<tr>
<td>July 31</td>
<td>No class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>August 1</td>
<td>One sample hypothesis tests using $Z$ and $t$ distributions</td>
<td>Ch. 10: 10.1–10.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Two sample hypothesis tests using $Z$ and $t$ distributions</td>
<td>Ch. 11: 11.1, 11.2</td>
<td></td>
</tr>
<tr>
<td>August 2</td>
<td>hypothesis tests (cont.)</td>
<td>Ch. 11: 11.1, 11.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Estimation and confidence intervals: Use and interpretation</td>
<td>Ch. 9: 9.1–9.3</td>
<td></td>
</tr>
<tr>
<td>August 7</td>
<td>Nonparametric tests</td>
<td>Ch. 13: 13.1-13.5</td>
<td>HW 4</td>
</tr>
<tr>
<td>August 8</td>
<td>Proportions: Estimation and hypothesis tests</td>
<td>Ch. 14: 14.1–14.4, 14.6</td>
<td></td>
</tr>
<tr>
<td>August 9</td>
<td>Chi-square ($\chi^2$) distribution and $\chi^2$ tests</td>
<td>Ch. 15: 15.1</td>
<td></td>
</tr>
<tr>
<td>August 14</td>
<td>Power and sample size calculation</td>
<td>Ch. 10: 10.5–10.7</td>
<td>HW 5</td>
</tr>
<tr>
<td>August 15</td>
<td>Simple linear regression</td>
<td>Ch. 18: 18.1, 18.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multiple linear regression</td>
<td>Ch. 19: 19.1</td>
<td></td>
</tr>
<tr>
<td>August 16</td>
<td>Logistic regression</td>
<td>Ch. 20: 20.1, 20.2, 20.3</td>
<td></td>
</tr>
<tr>
<td>August 21</td>
<td>Review and “Q &amp; A” Session</td>
<td></td>
<td></td>
</tr>
<tr>
<td>August 22</td>
<td>FINAL EXAM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>August 24</td>
<td>final homework set due</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Program in Public Health Syllabus Checklist

As course director, please be sure to use this form as a checklist to ensure proper submission of your syllabus for approval by the Program in Public Health (PPH) Curriculum Committee. Failure to do so can result in the return of the syllabus without approval. Please check off the appropriate responses and submit an electronic copy along with your syllabus.

Instructor: Juned Siddique

Course Name and Number: Introduction to Biostatistics, EpiBio 302
☒ I have read the “Template for Northwestern University PPH syllabi” document

I have included the following in my course syllabus:
☒ Course Number
☒ Name of Class
☒ Quarter Year
☒ Day and Time
☒ Location
☒ Instructor’s name
☒ Instructor’s title
☒ Email
☒ Phone
☒ Office hours (by appointment okay)
☒ Course Description
☒ Learning Objectives that are observable and measurable
☒ Texts and Readings
☒ Class assignments, projects, quizzes, tests
☒ Student Evaluation Components (well defined and objective)
☒ Course evaluation statement
☒ Academic Integrity statement
☒ Class schedule

My course meets the following Program in Public Health Course Competencies (check all that apply):
☒ Critically evaluate epidemiologic, prevention and health promotion, clinical outcomes, and health services research studies.
☐ Formulate a testable hypothesis relevant to public health practice and select and implement appropriate methods to test the hypothesis in an ethically appropriate fashion with contemporary information and computing resources.

☐ Locate and interpret vital statistics and other population-based data.

☐ Identify population needs for primary, secondary, and tertiary prevention and describe population-based, organizational, and individual behavioral change approaches designed to restore, improve, and maintain health.

☐ Identify challenges and opportunities that the economics and organization of health services create for maintaining and improving the public’s health.

☐ Describe the differences and overlap between clinical medicine and public health, identify ways in which the two disciplines can work in synergy, and recognize the advantages of interdisciplinary teamwork in achieving health objectives.

☐ Develop advocacy strategies for public policies that advance health goals.

☐ Demonstrate leadership potential as exemplified by effective writing, public presentation, and teaching.

☐ Describe the history and traditions of public health and their relevance to current and future practice.