Course Title: Applied Regression Analysis for Public Health Studies

Course Number: BIST 0551J

Course Location: Lectures in RM 1B, School of Public Health, Piscataway, NJ

Course Date & Time: Tuesday, 6:10 – 9:00 PM

Course Instructor: Sinae Kim, PhD., Assistant Professor
Biostatistics, Rutgers School of Public Health
sinae.kim@rutgers.edu & (732) 235-8816

Office Hours: Tuesday, 4:30 - 5:30PM, Swing Office A (1st floor of SPH)

Course Assistant: NA

Regression Analysis and Other Multivariable Methods, 5th Ed.
Duxbury Publisher [4th Edition is also acceptable]

Additional/Supplemental Readings/Resources:

5th Ed. Prentice Hall, NJ

Course Description: This course introduces students in graduate programs in public health to
regression analyses methods. The primary topics are simple, multiple linear regression models,
including analysis of covariance (ANCOVA), model diagnostics and model building. Logistic
regression for binary outcome will also be introduced. The emphasis will be interpretation and
applications. Students will learn how to use SAS for implementing regression methods.

Selected Department Competencies Addressed: The competencies addressed in this course
for the MPH for the Department of Biostatistics include:

- Integrate relevant scientific background to design experimental and observational
  studies in biomedical, clinical and public health research;
- Use statistical computer packages to organize, analyze and report collected data;
- Review and critique statistical methods and interpretations presented in published
  research studies, presentations or reports; and
- Communicate the results of statistical studies both in writing and orally to investigators
  and lay community members.

Please visit the Department webpages on the School of Public Health’s website at http://sph.rutgers.edu/
for additional competencies addressed by this course for other degrees and departments.

Course Objectives: By the completion of this course, students will be able to:
• Understand the fundamental concepts and assumptions of regression;
• Develop the ability to apply these concepts correctly using statistical software; and
• Develop the ability to interpret the results of an analysis properly.

Course Requirements and Grading:

• Course evaluation
  1. Midterm Examination (in class, closed book) 25 %
  2. Final Examination (in class, closed book) 30 %
  3. Homework (6 assignments) 20 %
  4. Class exercise (participation) 10 %
  5. Data analysis report 15 %
  Total: 100%

• Data analysis report: A group of students (2 per group) will complete a full analysis (linear regression) of a set of data that contains a continuous measured response and multiple covariates (at least 7 covariates, any type) and write individual report. Each group provides a set of data, which should NOT be used for any group member’s fieldwork. The emphasis should be on the statistical methodology and how it is applied. A plan will be due on October 17th. The plan should include a full description of the data set and variables to be used, and a set of clearly defined hypotheses. The final report is due on Dec. 12th.

• Grading policy
  1. Homework will be collected at the beginning of lecture on due date. Unless notifying the instructor beforehand, later submission of homework will NOT be graded.
  2. On all homework assignments/problem sets, students are encouraged to discuss with one another, but work should be carried out and written up independently. If any two identical write-ups are found, both homework assignments are considered failed.
  3. It is the students’ responsibility to make their papers legible. Unreadable work will NOT be graded.
  4. The students are asked to answer each question as accurately and concisely as possible. If it is necessary to attach the computer output with the homework assignment, ONLY the “essential” segments are required. DO NOT SUBMIT the complete output section or the log file. Otherwise, 50% of the points will be taken away.
**Course Schedule:** This table provides a general plan for the course; some deviations may be necessary.

<table>
<thead>
<tr>
<th>Date</th>
<th>Week</th>
<th>Session 1: Simple Linear Regression (one covariate)</th>
<th>Date</th>
<th>Week</th>
<th>Session 2: Multiple Linear Regression (multiple covariates)</th>
</tr>
</thead>
<tbody>
<tr>
<td>09/05</td>
<td>1</td>
<td>Course overview; review on intro biostatistics; measuring association; correlation coefficient (chap 4, 6)</td>
<td>10/10</td>
<td>6</td>
<td>Introduction of multiple linear regression (MLR); interpretation and point estimation (chap 8); Discussion on HW3; <strong>HW4 assignment</strong> will be posted</td>
</tr>
<tr>
<td>09/12</td>
<td>2</td>
<td>Introduction to simple linear regression with model assumptions and interpretation; regression parameter estimation and inference including confidence interval &amp; hypothesis testing; applications; (chap 5, 7); HW1 assignment</td>
<td>10/17</td>
<td>7</td>
<td>Inference procedure in MLR including confidence interval and hypothesis testing for a single variable; categorical covariates in MLR (chap 9, 10, 12); HW4 Due; Data analysis plan DUE</td>
</tr>
<tr>
<td>09/19</td>
<td>3</td>
<td>Continued in estimation of regression parameters; prediction and predictive interval; assessing model fit; (chap 5, 7) HW1 DUE; <strong>HW2 assignment</strong> will be posted</td>
<td>10/24</td>
<td>8</td>
<td>Introduction to analysis of covariance (ANCOVA), confounding factors and interaction (chap 11, 13 ); discussion on HW4;</td>
</tr>
<tr>
<td>09/26</td>
<td>4</td>
<td>Residual analysis to check model assumptions; (chap 14.1 – 14.4); discussion on HW1; HW2 DUE; <strong>HW3 assignment</strong> will be posted</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/03</td>
<td>5</td>
<td>Continued in model diagnostics (outlier; leverage point; influential point) (chap 14.1-14.4); Discussion on HW2 HW3 Due</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Subject</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/31</td>
<td>Review for midterm exam</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11/07</td>
<td>Testing for multiple effects using F test (chap 9, 10); discussion on midterm exam;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11/14</td>
<td>Model building and checking (chap 14, 16); discussion on data analysis project;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>HW5 assignment</strong> will be posted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Session 3:</strong> Logistic Regression Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11/21</td>
<td>No class (THURSDAY classes meet)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11/28</td>
<td>Review on Binomial distribution; introduction to logistic regression models; maximum likelihood estimation (chap 21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>HW5 DUE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>HW6 assignment</strong> will be posted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12/05</td>
<td>Parameter estimation and inference; goodness-of-fit test; model checking; examples; introduction of ungrouped data (chap 22); discussion on HW5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12/12</td>
<td>Logistic regression models II: ungrouped data; Introduction of case-control studies (chap 22); review for final exam; discussion on HW6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>HW6 DUE</strong>; Data analysis report DUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12/19</td>
<td>Final Exam</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**School of Public Health Honor Code:** The School of Public Health Honor Code is found in the student bulletin (sph.rutgers.edu/academics/catalog/index.html). Each student bears a fundamental responsibility for maintaining academic integrity and intellectual honesty in his or her graduate work. For example, all students are expected to observe the generally accepted principles of scholarly work, to submit their own rather than another’s work, to refrain from falsifying data, and to refrain from receiving and/or giving aid on examinations or other assigned work requiring independent effort. In submitting written material, the writer takes full responsibility for the work as a whole and implies that, except as properly noted by use of quotation marks, footnotes, etc., both the ideas and the works used are his or her own. In addition to maintaining personal academic integrity, each student is expected to contribute to the academic integrity of the school community by not facilitating inappropriate use of her/his own work by others and by reporting acts of academic dishonesty by others to an appropriate school authority. It should be clearly understood that plagiarism, cheating, or other forms of academic dishonesty will not be tolerated and can lead to sanctions up to and including separation from the Rutgers School of Public Health.
Policy Concerning Use of Recording Devices and Other Electronic Communications Systems:
When personally owned communication/recording devices are used by students to record lectures and/or classroom lessons, such use must be authorized by the faculty member or instructor who must give either oral or written permission prior to the start of the semester and identify restrictions, if any, on the use of mobile communications or recording devices.