

BIOSTAT 650
Applied Statistics I: Linear Regression
Fall 2008

Instructor: Sinae Kim, Ph.D., Assistant professor of Biostatistics
M4065, SPH II
sinae@umich.edu

Text (required): Montgomery, D.C., Peck, E.A. and Vining, G.G. (2006)
Introduction to Linear Regression Analysis, 4th Edition. Wiley, New York

Course Web-site: ctools.umich.edu; login using your unique name and password,
then you will see "BIOSTAT650 001 F08" on the main menu,
All class notes at the "Resources" directory.

Pre-/Co-requisite:
BIOSTAT 601, or similar masters level mathematical statistics course.
Familiarity with statistical methods like estimation (point, interval) and hypothesis testing,
and also with calculus, and matrix algebra is assumed.

Computing : SAS (various procedure); Calculator required for some of classes and exams.

Lectures : Monday & Wednesday 1:00 - 3:00 PM (Michigan Time)
M1112, SPH-II
No classes on Mon-06-Oct.(SPH symposium) and Mon-20-Oct.(Fall break).
Class ends on Mon-08-Dec.

Office Hours : Tuesday 3:00 - 5:00 PM; Wednesday 9 - 10 AM;

Grading : Homework 20% (approximately 8 homeworks)
Midterms 50% = $2 \times 25\%$ (Tentative schedule: Mon-13-Oct and Wed-19-Nov.)
Final exam 30% (University schedule: Thurs-11-Dec., 4-6 PM)

Topics (tentative list):

1. Before Exam 1
 - (a) Introduction/Review
 - (b) Simple linear regression
 - i. Least-squares estimation (LSE)/Hypothesis testing: Slope and Intercept.
 - ii. Interval estimation: confidence interval
 - iii. Prediction of new observations: predicted value & prediction interval
 - (c) Multiple linear regression
 - i. LSE/Hypothesis testing: Regression coefficients
 - ii. General linear hypothesis testing
 - iii. Interval estimation
 - iv. Prediction of new observations
2. Before Exam 2
 - (a) Indicator variables
 - i. Dummy variable; cell coding
 - (b) Residual diagnostics
 - i. Definition of residuals (5 different types of residuals)
 - ii. Residual plots
 - iii. Partial regression & partial residual plots
 - iv. Lack-of-Fit (LOF) test
 - (c) Transformations
 - i. Variance-stabilizing transformation (VST)
 - ii. Linearization
 - iii. Box-Cox transformation
 - iv. Generalized and Weighted least squares estimation
 - (d) Influence diagnostics
 - i. Leverage
 - ii. Measures of Influence: Cook's distance, DFFITS, DFBETAS, COVRATIO etc.
3. Before the end of class
 - (a) Variable selection
 - i. All possible regressions
 - ii. Forward, backward and stepwise selection
 - (b) Multicollinearity
 - i. Multicollinearity diagnostics: correlation matrix, variance inflation factors, eigenvalues

- ii. Dealing with multicollinearity
- (c) Model validation

Academic Integrity:

The faculty of the School of Public Health believes that the conduct of a student registered or taking courses in the School should be consistent with that of a professional person. Courtesy, honesty and respect should be shown by students toward faculty members, guest lecturers, administrative support staff and fellow students. Similarly, students should expect faculty to treat them fairly, showing respect for their ideas and opinions and striving to help them achieve maximum benefits from their experience in the School. Student academic misconduct refers to behavior that may include plagiarism, cheating, fabrication, falsification of records or official documents, intentional misuse of equipment or materials (including library materials), and aiding and abetting the perpetration of such acts. The preparation of reports, papers, and examinations, assigned on an individual basis, must represent each student's own effort. Reference sources should be indicated clearly. The use of assistance from other students or aids of any kind during a written examination, except when the use of aids such as electronic devices, books or notes has been approved by an instructor, is a violation of the standard of academic conduct.