

PUB HLTH 421
INTERMEDIATE BIOSTATISTICS – 1.0 credit
Winter Quarter 2008 (January 7 – March 17, 2008)
Time: Mondays & Wednesdays 6:00-7:30 p.m.
Location: Lurie Building, Gray Seminar Room, 303 E. Superior

Course instructor (office hours by appointment):

Alan R. Dyer, Ph.D.
Professor
Department of Preventive Medicine
(312) 908-7919
adyer@northwestern.edu

I. Course Description

This course builds upon the material learned in Introduction to Biostatistics and Intermediate Epidemiology. Specifically, the course will focus on multivariate methods of analysis for epidemiologic and clinical studies including correlation, linear regression, logistic regression, and Cox proportional hazards regression.

II. Prerequisites

Required:

- a) Introduction to Epidemiology – PH304 (or Medical Decision Making II)
- b) Introduction to Biostatistics – PH302

Recommended:

- a) Intermediate Epidemiology or concurrent registration in Intermediate Epidemiology – PH422

The student **MUST** receive permission from the instructor if the two required courses listed above have not been completed.

III. Course Objectives

This course is designed to familiarize students with additional statistical techniques and multivariate methods of analysis for epidemiological and clinical studies. Upon completion of this course students should be able to:

- Describe the difference between a correlation and regression analysis.
- Describe the assumptions and mechanics for estimating parameters in the simple and multiple linear regression models.
- Use Stata to perform a simple linear regression or multiple linear regression analysis, and interpret the output, including regression parameter estimates, hypothesis tests, confidence intervals, statistics which quantify the fit of the model, and potential confounding.
- Use graphical procedures in Stata for assessing whether the assumptions underlying the simple or

multiple linear regression models are correct. Describe ways to remedy the problem when the assumptions appear incorrect.

- Use graphical procedures for assessing whether there are outliers in a linear regression analysis and whether any of the observations exerts considerably more influence than the other observations in determining estimates of regression coefficients.
- Use Stata to perform an analysis of variance (ANOVA) and covariance (ANCOVA), and interpret the output, including sums of squares and F-tests.
- Use Stata for multiple comparisons and interpret results for ANOVA and ANCOVA.
- Describe the logistic regression model, its key assumptions, and their implications.
- State the relationships between odds ratios and logistic regression coefficients.
- Use Stata to fit a logistic regression model to continuous and categorical predictors and to obtain odds ratios, and interpret the output, including regression parameter estimates, hypothesis tests, confidence intervals, and statistics which quantify the fit of the model.
- Derive the odds ratio between two groups defined by their predictor values.
- Calculate and interpret the likelihood ratio test comparing two nested logistic models.
- Use Stata to obtain Kaplan-Meier survival curves, and interpret results.
- Use Stata to perform a log-rank test and interpret results.
- Define right censoring, hazard function, and proportional hazards.
- Use Stata to fit a Cox proportional hazards regression model to continuous and categorical predictors and to obtain hazard ratios, and interpret the output, including regression parameter estimates, hypothesis tests, and confidence intervals.
- Derive the hazard ratio between two groups defined by their predictor values.
- Calculate and interpret the likelihood ratio test comparing two nested Cox models.
- Detect non-proportional hazards using log-minus-log plots.
- Describe situations in which multiple linear regression analysis, logistic regression analysis, or survival analysis is needed.
- Detect and model interaction in these multi-predictor models.
- Use dummy variables to model categorical predictors and interpret results in each of these models.
- Interpret articles in the medical literature that use multiple linear regression, multiple logistic regression, or Cox proportional hazards regression.

IV. Teaching Format

New material will be presented in informal lectures. Homework assignments will be handed in for evaluation. Students will be expected to read the assigned materials prior to the lecture or the discussion. Participation in group discussions is mandatory.

V. Student Evaluation

a) Homework and classroom participation - Homework assignments will count 20% towards the student evaluation. Homework assigned on a Monday or Wednesday is due on the following Monday. Answers to homework assignments will be posted on **Blackboard** on the following Tuesday. Missed homework assignments will not be accepted once answers are posted on **Blackboard**. Students are expected to come to class prepared and to take part in the classroom discussions.

b) Examinations - There will be a midterm examination (30%) and a comprehensive final examination (50%). The examinations may consist of multiple choice, short answer, and statistical computations or

interpretation of Stata output, and will be administered at the scheduled times. Make up examinations should be arranged in advance and will only be given under extenuating circumstances.

VI. Textbooks

Required: None

Optional*: Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE. Regression Methods in Biostatistics. Linear, Logistic, Survival and Repeated Measures Models. Springer, New York, NY, 2005.

*This book will be on reserve in the library.

VII. Required Software

Students must have access to Intercooled Stata or Stata/SE. It is highly recommended that students purchase one of these versions, if they do not already own one. Homework assignments require that students do analyses using Stata. Stata/SE is installed on 14 PCs in the Galter Learning Resources Center (LRC). Stata/SE is also installed on 3 PCs in the MPH Student's Room in Suite 1220 of 680 N. Lake Shore Drive (use the Guest account which has no password).

VIII. Blackboard

The course management system, Blackboard, will be used to transmit and share all course materials including the syllabus, assignments, and lecture presentations.

IX. Course Evaluation

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 to complete the evaluations before grades are submitted.

XI. Academic Integrity

Academic integrity at Northwestern University is based on a respect for individual achievement that lies at the heart of academic culture. Every faculty member and student, both graduate and undergraduate, belongs to a community of scholars where academic integrity is a fundamental commitment. The Programs in Public Health abides by the standards of academic conduct, procedures, and sanctions as set forth by The Graduate School at Northwestern University. Students are responsible for knowledge of the information provided by The Graduate School on their Web page at <http://www.tgs.northwestern.edu/studentsvcs/ethics/>.

Intermediate Biostatistics

Session	Date	Format	Topic	Readings
1	Mon 1/07	Lecture	Introduction and Overview	Passed out in class
2	Wed 1/09	Lecture	Simple Linear Regression	Blackboard: Simple Linear Regression
3	Mon 1/14	Lecture	Diagnostics in Simple Linear Regression	Blackboard: Diagnostics in Simple Linear Regression
4	Wed 1/16	Lecture	Diagnostics in Simple Linear Regression-II Introduction to Multiple Linear Regression	Blackboard: Diagnostics in Simple Linear Regression-II; Introduction to Multiple Linear Regression
	Mon 1/21	No class		
5	Wed 1/23	Lecture	Multiple Linear Regression	Blackboard: Multiple Linear Regression
6	Mon 1/28	Lecture	Diagnostics in Multiple Linear Regression	Blackboard: Diagnostics in Multiple Linear Regression
7	Wed 1/30	Lecture	Additional Topics in Multiple Linear Regression	Blackboard: Additional Topics in Multiple Linear Regression
8	Mon 2/4	Paper discussion	Multiple Linear Regression in the Medical Literature	To be determined
9	Wed 2/6	Lecture	Analysis of Variance and Covariance	Blackboard: Analysis of Variance and Covariance
10	Mon 2/11	EXAM	Midterm examination	

11	Wed 2/13	Lecture	Introduction to Logistic Regression	Blackboard: Introduction to Logistic Regression
12	Mon 2/18	Lecture	Multiple Logistic Regression - I	Blackboard: Multiple Logistic Regression – I
13	Wed 2/20	Lecture	Multiple Logistic Regression - II	Blackboard: Multiple Logistic Regression – II
14	Mon 2/25	Paper discussion	Multiple Logistic Regression in the Medical Literature	To be determined
15	Wed 2/27	Lecture	Introduction to Person-Time Data	Blackboard: Introduction to Person-Time Data
16	Mon 3/03	Lecture	Introduction to Cox Proportional Hazards Regression	Blackboard: Introduction to Cox Proportional Hazards Regression
17	Wed 3/05	Lecture	Cox Proportional Hazards Regression	Blackboard: Cox Proportional Hazards Regression
18	Mon 3/10	Paper discussion	Cox Proportional Hazards Regression in the Medical Literature	To be determined
19	Mon 3/12	Lecture	Course Review	Blackboard: Course Summary
20	Mon 3/17	EXAM	Comprehensive Final Examination	

09/14/08