Biostatistics 406
Applied Multivariate Biostatistics

Instructor: Professor A. A. Afifi

Course Topics: Various topics in applied multivariate analysis including multiple linear regression, logistic regression, log-linear models, principal components and factor analysis and survival analysis.

Prerequisites: Biostatistics 100B or equivalent.

Learning Objectives:

COURSE LEARNING OBJECTIVES	HOW THESE LEARNING OBJECTIVES ALIGN WITH COMPETENCIES FOR SPECIFIC DEGREE PROGRAMS				
	Competencies for MPH in Biostatistics	Competencies for MS in Biostatistics	Competencies for DrPH in Biostatistics	Competencies for PhD in CHS	Competencies for PhD in Epidemiology
1. Fit and compare different models to analyze health science data.	G4. Consult with public health professionals and researchers helping them design research studies (using statistically rigorous methods for sample size determination and power) and analyzing data obtained from such designs. G7. Describe preferred methodological alternatives to commonly used statistical methods when assumptions are not met.	A6. Conduct appropriate statistical analyses of study data and interpret the results. A7. Effectively communicate the assumptions and results of analyses through oral and written communications to the collaborative team.	B6. Conduct appropriate statistical analyses of study data and interpret the results. B7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team.	4. Demonstrate expertise in advanced research methods (including research design and implementation, data analysis, and statistics) in the social and behavioral sciences and apply these methods to conduct hypothesis-testing and/or hypothesis-generating research in the student's own area of research.	7. Demonstrate expertise in advanced epidemiologic research methods (including research design and implementation, data analysis and statistics) and apply these methods to conduct hypothesis-testing research in the student's own area of research

COURSE LEARNING OBJECTIVES	HOW THESE LEARNING OBJECTIVES ALIGN WITH COMPETENCIES FOR SPECIFIC DEGREE PROGRAMS				
	Competencies for MPH in Biostatistics	Competencies for MS in Biostatistics	Competencies for DrPH in Biostatistics	Competencies for PhD in CHS	Competencies for PhD in Epidemiology
2. Use various SAS procedures such as REG, LOGISTIC, CATMOD, LIFETEST, PHREG, PRINCOMP, and FACTOR.	G3. Develop analytical and computational skills for the management, modeling and analysis of public health datasets with several variables that may be dependent on one another using statistically rigorous methods and models.	A6. Conduct appropriate statistical analyses of study data and interpret the results. A7. Effectively communicate the assumptions and results of analyses through oral and written communications to the collaborative team.	B6. Conduct appropriate statistical analyses of study data and interpret the results. B7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team.	4. Demonstrate expertise in advanced research methods (including research design and implementation, data analysis, and statistics) in the social and behavioral sciences and apply these methods to conduct hypothesis-testing and/or hypothesis-generating research in the student's own area of research.	7. Demonstrate expertise in advanced epidemiologic research methods (including research design and implementation, data analysis and statistics) and apply these methods to conduct hypothesis-testing research in the student's own area of research
3. Perform a multiple linear, logistic or survival regression and correctly interpret the output, including the coefficients and their confidence intervals and related tests of hypotheses.	G2. Develop analytical skills and obtain broad insights involving the design and analysis of experiments to understand and model the dependence between different variables (e.g. regression), handle missing or incomplete data, and carry out rigorous statistical modeling for data obtained from a variety of public health study designs.	A6. Conduct appropriate statistical analyses of study data and interpret the results. A7. Effectively communicate the assumptions and results of analyses through oral and written communications to the collaborative team.	B6. Conduct appropriate statistical analyses of study data and interpret the results. B7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team.	4. Demonstrate expertise in advanced research methods (including research design and implementation, data analysis, and statistics) in the social and behavioral sciences and apply these methods to conduct hypothesis-testing and/or hypothesis-generating research in the student's own area of research.	7. Demonstrate expertise in advanced epidemiologic research methods (including research design and implementation, data analysis and statistics) and apply these methods to conduct hypothesis-testing research in the student's own area of research

COURSE LEARNING OBJECTIVES	HOW THESE LEARNING OBJECTIVES ALIGN WITH COMPETENCIES FOR SPECIFIC DEGREE PROGRAMS				
	Competencies for MPH in Biostatistics	Competencies for MS in Biostatistics	Competencies for DrPH in Biostatistics	Competencies for PhD in CHS	Competencies for PhD in Epidemiology
4. Examine plots and other analyses for inconsistencies between the fitted model and patterns in the data and for outliers and high leverage observations.	G2. Develop analytical skills and obtain broad insights involving the design and analysis of experiments to understand and model the dependence between different variables (e.g. regression), handle missing or incomplete data, and carry out rigorous statistical modeling for data obtained from a variety of public health study designs..	A6. Conduct appropriate statistical analyses of study data and interpret the results. A7. Effectively communicate the assumptions and results of analyses through oral and written communications to the collaborative team.	B6. Conduct appropriate statistical analyses of study data and interpret the results. B7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team.	4. Demonstrate expertise in advanced research methods (including research design and implementation, data analysis, and statistics) in the social and behavioral sciences and apply these methods to conduct hypothesis-testing and/or hypothesis-generating research in the student's own area of research.	7. Demonstrate expertise in advanced epidemiologic research methods (including research design and implementation, data analysis and statistics) and apply these methods to conduct hypothesis-testing research in the student's own area of research
5. Select the best model for the data based on model fit statistics and variable selection methods.	G2. Develop analytical skills and obtain broad insights involving the design and analysis of experiments to understand and model the dependence between different variables (e.g. regression), handle missing or incomplete data, and carry out rigorous statistical modeling for data obtained from a variety of public health study designs.	A6. Conduct appropriate statistical analyses of study data and interpret the results. A7. Effectively communicate the assumptions and results of analyses through oral and written communications to the collaborative team.	B6. Conduct appropriate statistical analyses of study data and interpret the results. B7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team.	4. Demonstrate expertise in advanced research methods (including research design and implementation, data analysis, and statistics) in the social and behavioral sciences and apply these methods to conduct hypothesis-testing and/or hypothesis-generating research in the student's own area of research.	7. Demonstrate expertise in advanced epidemiologic research methods (including research design and implementation, data analysis and statistics) and apply these methods to conduct hypothesis-testing research in the student's own area of research

COURSE LEARNING OBJECTIVES	HOW THESE LEARNING OBJECTIVES ALIGN WITH COMPETENCIES FOR SPECIFIC DEGREE PROGRAMS				
	Competencies for MPH in Biostatistics	Competencies for MS in Biostatistics	Competencies for DrPH in Biostatistics	Competencies for PhD in CHS	Competencies for PhD in Epidemiology
6. Perform a principal components, factor or log-linear analysis to explore underlying associations among a group of variables.	G2. Develop analytical skills and obtain broad insights involving the design and analysis of experiments to understand and model the dependence between different variables (e.g. regression), handle missing or incomplete data, and carry out rigorous statistical modeling for data obtained from a variety of public health study designs.	A6. Conduct appropriate statistical analyses of study data and interpret the results. A7. Effectively communicate the assumptions and results of analyses through oral and written communications to the collaborative team.	B6. Conduct appropriate statistical analyses of study data and interpret the results. B7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team.	4. Demonstrate expertise in advanced research methods (including research design and implementation, data analysis, and statistics) in the social and behavioral sciences and apply these methods to conduct hypothesis-testing and/or hypothesis-generating research in the student's own area of research.	7. Demonstrate expertise in advanced epidemiologic research methods (including research design and implementation, data analysis and statistics) and apply these methods to conduct hypothesis-testing research in the student's own area of research

COURSE LEARNING OBJECTIVES	HOW THESE LEARNING OBJECTIVES ALIGN WITH COMPETENCIES FOR SPECIFIC DEGREE PROGRAMS				
	Competencies for MPH in Biostatistics	Competencies for MS in Biostatistics	Competencies for DrPH in Biostatistics	Competencies for PhD in CHS	Competencies for PhD in Epidemiology
7. Participate in a research team to formulate scientific questions, analyze them using various methods and write a report describing the analysis and conclusions.	G8. Develop written and oral presentations based on statistical analyses for public health professionals as well as lay audiences.	A1. Collaborate with researchers to formulate the aims of a public health research project. A2. Formulate a public health question in statistical terms. B2. Effectively communicate statistical concepts and reasoning to public health collaborators. B3. Learn to write and disseminate substantive field publications and communicate the statistical portion of the methodology to a substantive field audience. B5. Be able to articulate interdisciplinary approaches to solving public health problems.	B1. Collaborate with researchers to formulate the aims of a public health research project. B2. Formulate a public health question in statistical terms. C4. Effectively communicate statistical concepts and reasoning to public health collaborators. C5. Learn to write and disseminate substantive field publications and communicate the statistical portion of the methodology to a substantive field audience. D5. Understand and be able to effectively communicate the public health significance of the problems being addressed.	5. Formulate a research question on an important public health topic, design a rigorous and original empirical study to answer it, conduct that study, interpret the results, and draw conclusions. 8. Disseminate research findings, including prepare a scientific article suitable for a refereed journal based on an original research project, submit the article for publication, respond to critiques of journal reviewers, and take appropriate action in response to a rejection.	8. Formulate a research question on an important epidemiologic topic, design a rigorous and original empirical study to answer it, conduct that study, interpret the results, and draw conclusions

Homework Assignments: You will be required to hand in several assignments throughout the quarter based on a combination of problems in the textbook and data analyses from the computer lab.

Midterm and Final Examinations: There will be one midterm and a final examination held in class.

Group Final Paper: You are required to do a Group Final Paper. The group final paper is to be co-authored by at least 3 (max. 4) enrolled students of the class (no auditors allowed). You will need to plan early. The final paper should summarize and extend the data analyses from the homework assignments in a more formal manner and is considered a "mini-publication". All co-authors will receive the same grade for the Group Final Paper. Additional information on the specific requirements and the deadline for submitting it will be given in class.

Grading: Grades will be based on the homework assignments (20%), the midterm (20%), the final (20%) and the group final paper (40%).

Data Sets: The data sets that are used for homework assignments will be available for downloading on the internet. Some of the analyses in the course will use data sets described in the course text and others will be focused on a large data set supplied by the instructor. Supplementary materials describing this data set will be available on the course web site.

Computing: In the lab sessions you will be using the SAS language. You should have at least some familiarity with SAS; if not be prepared to spend some extra time on the computer in the first couple of weeks.

Course Text:

The textbook for the class is *Practical Multivariate Analysis*, 5th Ed., by A. A. Afifi, S. J. May and V. A. Clark, Chapman & Hall, New York: 2012. It is available in the Health Sciences Bookstore.

Internet: There is a web page for this course at //afifi.bol.ucla.edu/biostat406/

Tentative schedule

Week	Topic	Textbook Chapters
1	Introduction and preview	1 - 5
	Review simple linear regression	6
2	Review multiple linear regression	7
	Variable selection	8
3	Other linear regression topics	9
	Survival analysis	13
4	Survival analysis, continued	13
5	Logistic regression	12
	Midterm	
6	Logistic regression, continued	12
	Log-linear models	17
7	Principal components analysis	14
8	Factor analysis	15
9	Missing values analysis	9
10	Cluster analysis	16
	Review	