BIOS 663: Intermediate Linear Models

Texts:

- Muller and Fetterman, *Regression and ANOVA: An Integrated Approach Using SAS Software* (required)
- SAS/STAT User's Guide (in particular for PROC REG and PROC GLM; available on Blackboard web site) (required)
- Kleinbaum, Kupper, Muller, and Nizam, *Applied Regression Analysis and Other Multivariable Methods* (recommended)
- Harrell, *Regression Modeling Strategies* (recommended)

BIOS663 offers a matrix-based treatment of regression, one-way and two-way ANOVA, and ANCOVA, with applications in the health sciences. Emphasis is placed on the general linear model and hypothesis, diagnostics, and model building, as well as on interpretation and communication of statistical results. The course begins with a review of matrix algebra and concludes with introductions to longitudinal data analysis and generalized linear models.

In BIOS663, students will

1. learn to analyze and interpret univariate linear models,
2. gain a basic understanding of the associated theory of linear models in order to know when not to apply the methods and how to extend the theory to non-standard situations,
3. gain exposure to more advanced models, including generalized linear models and mixed effects models, and
4. learn to communicate statistical results to subject-area collaborators (including public health professionals, physicians, and other scientists and researchers) in non-statistical language so that the results can be understood.

Course Prerequisites: The following prerequisite skills are assumed.

- Working knowledge of calculus and matrix algebra (through MATH 547)
- Working knowledge of probability and inference (through BIOS 550/660)
- General applied statistics knowledge (at the level of BIOS 662)
- Working knowledge of SAS (at the level of BIOS 511)

Students are responsible for ensuring they have the necessary prerequisites. Violations of assumptions will likely lead to frustration on the part of the student! BIOS 663 is primarily for students in the biostatistics department, though quantitative-minded students from other departments are welcome (if you are such a student, see me if you have questions about your mathematical background). The department also offers two other courses in linear models: BIOS 762 is an advanced theoretical treatment of linear models, and BIOS 545 is a basic applied course in linear models. Students without the necessary prerequisites (MATH 547, BIOS 511, BIOS 550 or BIOS 660, BIOS 662) should consider other courses in the BIOS department or in the school.

Copies of overhead transparencies used in class will be available online at

`http://blackboard.unc.edu`

. These notes cover some material not contained in the texts and do not cover all material in the texts so that the assigned readings are very important. You are encouraged to print the course notes and bring them with you to class.

Graded Work : Grades will be determined as follows:

- 10% in-class problems/participation
- 10% homework assignments
- 15% data analysis project
- 30% midterm exam
- 35% final exam

Graduate courses in the School of Public Health use the following grading system.

- **H**: Clear excellence
- **P**: Entirely satisfactory
- **L**: Low passing
- **F**: Fail

All assignments are cumulative. The SPH grading system is designed so that the mode of the grading distribution is **P**.

Descriptions of Graded Work:

- **Homework and In-Class Problems** Homework and in-class problems are designed to ensure that the text and readings have been read and understood and that material from the text and course notes has been mastered. Biweekly homework assignments will be given, and students are encouraged to talk about ideas and approaches to problems in groups, though students should "write up" assignments independently. In addition, problems will be given during class time. These problems will serve as an indicator of your strengths and weaknesses before the exams. Missed problems may not be made up, though up to 2 problem scores may be replaced (either due to absence or poor performance) with your score on the midterm exam.

- **Data Analysis Project** The project will consist of applying concepts learned in this class to data of your own choosing. Students may work in groups of two or three. Your project will be presented as a poster. More details on the data analysis project will be available on a project handout.

- **Exams** Each exam will consist of in-class and take-home portions. The exams are designed to test your mastery of the material presented, your ability to conduct statistical

analyses on your own, and your ability to interpret the results of analyses in language that subject-area investigators can easily understand. Students will *not* be allowed to work together while take-home exams are being given.

The Student Honor Code: Students are encouraged to work together on homework, but copying someone else's work *always* creates an honor code violation. Students are never allowed to discuss or work together on exams while they are being given. Expulsion from the university is possible if the honor code is violated, and receiving 0% on the assignment in question is a certainty.

Advice: This course will not be easy. I expect students to attend and participate in lectures, to read and understand the book, and to take homeworks, exams, and the class project seriously. I cannot overemphasize the importance of keeping up with the course material each day. Students who succeed in graduate courses typically work 6-8 hours per week outside of the classroom on each course. You may find that study groups are very useful, not only when preparing for exams, but also throughout the semester in preparing homeworks or discussing topics.