NC STATE UNIVERSITY

ST 537
APPLIED MULTIVARIATE AND LONGITUDINAL DATA ANALYSIS

ST 437/537: Applied Multivariate and Longitudinal Data Analysis

Section 001
Spring 2019
3 Credit Hours
Course website: https://maityst537.wordpress.ncsu.edu/

Course Description

An introduction to use of statistical methods for analyzing multivariate data (multiple variables or traits measured for the same individual) and longitudinal data (same variable or trait repeatedly measured on individuals over time) collected in experiments and surveys. Topics covered include multivariate analysis of variance, discriminant analysis, principal components analysis, factor analysis, covariance modeling, and mixed effects models such as growth curves and random coefficient models. Emphasis is on the use of a computer to perform statistical analysis of multivariate and longitudinal data. (Previously, students had to take ST 731 and ST 732 to cover these topics. This course combines topics from those courses and be aimed at students across campus and Statistics Masters students.)

Learning Outcomes

Students will learn basic techniques and results of multivariate and longitudinal data analysis. Upon completion of the course students will be able to:

- Distinguish between multivariate data and longitudinal data and adequately suggest methods for analysis
- Compare multivariate techniques and longitudinal methods and identify their advantages/limitations
- Use a computer (use R) to carry out analysis of multivariate and longitudinal data
- Correctly interpret the computer output in these situations

Course Structure

The course meets for lectures twice per week. Students are required to complete homework (about ten homework assignments), two mid-term exams, and a final R project (done in small groups).

Course Policies

Unexcused late homework will not be accepted. The final homework average will be computed after dropping the two lowest grades. You are permitted to work together on the homework sets, but each student is responsible for their final write-up of each assignment. When asked to solve problems using a computer, please provide well commented and neatly written computer code, relevant computer output, highlight all relevant results, and give an appropriate discussion. Examinations will be closed book and closed notes. A final R project will replace the final exam. Students may bring calculators to all tests, in addition to pen/pencil and scratch papers. No cell phones or other electronic devices should be in sight or used in any way during exams. Course policies are subject to change.

Instructors

Arnab Maity – Instructor
Email: amaity@ncsu.edu
Web: https://www.stat.ncsu.edu/people/maity/
Phone: 919-515-1937
Office location: SAS 5240
Office hours: Monday 3 – 4 pm

Zhou Lan – Teaching Assistant
Email: zlan@ncsu.edu
Office Location: 1101 SAS Hall (Statistics Tutorial center)
Office hours: Tuesday 8:30 – 10:00 am and Wednesday 8:30 – 10:00 am

Course Meetings

Lecture Days: Tu/Th
Time: 10:15am – 11:30am
Location: 1216 SAS Hall
This meeting is required.

Requisites and Restrictions

Prerequisites: ST 512 or 514 or 515 or 517.
Co-requisites: None
Restrictions: None

Course Materials

Textbooks: Materials will be taken from a variety of books. These books are **freely available from NCSU library** (see the links below).

- Applied Multivariate Statistics with R by Daniel Zelterman. New York: Springer
- An Introduction to Applied Multivariate Analysis with R by Brian Everitt and Torsten Hothorn. New York: Springer.
- Modeling Longitudinal Data by Robert E. Weiss. New York: Springer.
- Linear Mixed Models for Longitudinal Data by Geert Verbeke and Geert Molenberghs. New York: Springer.
- Applied Longitudinal Analysis by Fitzmaurice by G.M., Laird, N.M., and Ware, J.H. New York: Wiley (**on reserve**)

Expenses: None
Materials: R software available at https://www.r-project.org/ for free. *This material is required*. You may also want to download R studio, a free and useful IDE for R available at https://www.rstudio.com/

General Education Program (GEP) Information

GEP Category: This course does not fulfill a General Education Program category.
GEP Co-requisites: This course does not fulfill a General Education Program co-requisite.

Transportation

This course will not require students to provide their transportation. Non-scheduled class time for field trips or out-of-class activities is not required for this class.

Safety & Risk Assumptions

None.

Grading

The final numeric grade will be computed based on the following components. The exam and project deadlines are subject to change.

- **Homeworks (40%):** The final homework average will be computed after dropping the two lowest grades. As the lowest two scores are dropped, no late assignments are accepted.
- **Midterm Exam I (20%):** Tentatively scheduled on **March 07, 2019 (Thursday)** in class during lecture time. The examination will be closed book and closed notes. The student must contact the instructor in advance if s/he is likely to miss the scheduled midterm exam.
- **Midterm Exam II (20%):** Tentatively scheduled on **April 18, 2019 (Thursday)** in-class during lecture time. The examination will be closed book and closed notes. The student must contact the instructor in advance if s/he is likely to miss the scheduled midterm exam.
- **Final R project (20%):** There will be a larger R project (done in small groups of 3) toward the end of the semester. Students should submit a neatly typed project report along with R codes and output. All submitted code should be well commented. Your project **proposal is due on April 02, 2019**. Project **reports are due on April 25, 2019 (Thursday)**.

This Course uses Standard NCSU Letter Grading:

A+ ≥ 97 > A ≥ 93 > A– ≥ 90
B+ ≥ 87 > B ≥ 83 > B– ≥ 80
C+ ≥ 77 > C ≥ 73 > C– ≥ 70
D+ ≥ 67 > D ≥ 63 > D– ≥ 60 > F.

Requirements for Credit-Only (S/U) Grading

To receive a grade of S, students are required to take all exams and quizzes, complete all assignments, and earn a grade of C- or better. Conversion from letter grading to credit only (S/U) grading is subject to university deadlines. Refer to the Registration and Records calendar for deadlines related to grading. For more details refer to http://policies.ncsu.edu/regulation/reg-02-20-15.

Requirements for Auditors (AU)

Information about and requirements for auditing a course can be found at http://policies.ncsu.edu/regulation/reg-02-20-04. Auditors are expected to attend class regularly and submit homework on the same schedule as the other students. The final grade for

auditors (AU or NR) will be based on their final homework average (final homework grade will be calculated by dropping the two lowest grades). A final homework score of at least 70% is required for an AU.

Policies on Incomplete Grades

If an extended deadline is not authorized by the Graduate School, an unfinished incomplete grade will automatically change to an F after either (a) the end of the next regular semester in which the student is enrolled (not including summer sessions), or (b) by the end of 12 months if the student is not enrolled, whichever is shorter. Incompletes that change to F will count as an attempted course on transcripts. The burden of fulfilling an incomplete grade is the responsibility of the student. The university policy on incomplete grades is located at http://policies.ncsu.edu/regulation/reg-02-50-03

Additional information relative to incomplete grades for graduate students can be found in the Graduate Administrative Handbook in Section 3.18.F at http://www.fis.ncsu.edu/grad_publicns/handbook/

Late Assignments

Homework is due in class on the due date. No late assignments are accepted. If an emergency arises that prevents you from completing your work on time, please email the instructor as soon as possible so that arrangements can be made for you to keep up in the class.

Attendance Policy

For complete attendance and excused absence policies, please see http://policies.ncsu.edu/regulation/reg02-20-03. Students are expected to attend all lectures and exams.

Absences Policy: None.

Makeup Work Policy: No late HW submissions are accepted as two lowest scores will be dropped. Students who are unable to attend an exam for a legitimate unavoidable reason may take a make-up exam only if the student provides proper documentation of the delay and they can take the make-up exam in a very timely manner. If a make-up can't be taken then one of the midterm exams will be reweighted for the missing midterm exams. This may only be done with one of the exams. Students are required to take at least one of the two midterm exams on scheduled time otherwise a grade of F will be assigned.

Academic Integrity

Students are required to comply with the university policy on academic integrity found in the Code of Student Conduct found at http://policies.ncsu.edu/policy/pol-11-35-01

Academic Honesty: See http://policies.ncsu.edu/policy/pol-11-35-01 for a detailed explanation of academic honesty.

Honor Pledge: Your signature on any test or assignment indicates "I have neither given nor received unauthorized aid on this test or assignment."

While students are allowed to work in groups on homework, no one should copy directly from someone else's work (this includes present or past students). It is strongly urged that everyone work on their own as much as possible.

Electronically-Hosted Course Components

Students may be required to disclose personally identifiable information to other students in the course, via electronic tools like email or web-postings, where relevant to the course. Examples include online discussions of class topics and posting of student coursework. All students are expected to respect the privacy of each other by not sharing or using such information outside the course.

Electronically-hosted Components: https://maityst501.wordpress.ncsu.edu/

Accommodations for Disabilities

Reasonable accommodations will be made for students with verifiable disabilities. In order to take advantage of available accommodations, students must register with the Disability Services Office at Suite 2221, Student Health Center, Campus Box 7509, 919-515-7653. For more information on NC State's policy on working with students with disabilities, please see the Academic Accommodations for Students with Disabilities Regulation (REG 02.20.01).

Non-Discrimination Policy

NC State University provides equality of opportunity in education and employment for all students and employees. Accordingly, NC State affirms its commitment to maintain a work environment for all employees and an academic environment for all students that is free from all forms of discrimination. Discrimination based on race, color, religion, creed, sex, national origin, age, disability, veteran status, or sexual orientation is a violation of

state and federal law and/or NC State University policy and will not be tolerated. Harassment of any person (either in the form of quid pro quo or creation of a hostile environment) based on race, color, religion, creed, sex, national origin, age, disability, veteran status, or sexual orientation also is a violation of state and federal law and/or NC State University policy and will not be tolerated. Retaliation against any person who complains about discrimination is also prohibited. NC State's policies and regulations covering discrimination, harassment, and retaliation may be accessed at http://policies.ncsu.edu/policy/pol-04-25-05 or http://www.ncsu.edu/equal_op/. Any person who feels that he or she has been the subject of prohibited discrimination, harassment, or retaliation should contact the Office for Equal Opportunity (OEO) at 919-515-3148.

N.C. State University Policies, Regulations, and Rules (PRR)

Students are responsible for reviewing the PRRs which pertain to their course rights and responsibilities. These include http://policies.ncsu.edu/policy/pol-04-25-05 (Equal Opportunity and Non-Discrimination Policy Statement), http://oied.ncsu.edu/oied/policies.php (Office for Institutional Equity and Diversity), http://policies.ncsu.edu/policy/pol-11-35-01 (Code of Student Conduct), and http://policies.ncsu.edu/regulation/reg-02-50-03 (Grades and Grade Point Average).

Miscellany

- Attendance is expected at all lectures.
- Disputes about homework/exam grading must be brought to the instructor's attention within one week after the graded paper is returned.
- Students may discuss the homework problems with others. However, each student must submit their own independent write-up of the solutions. *Copying someone else's work—including* online *resources—is not acceptable and may result in disciplinary action*. The instructor is committed to upholding the university's policy on academic integrity, as described in the Code of Student Conduct. http://policies.ncsu.edu/policy/pol-11-35-01
- Students are responsible for reading, understanding, and adhering to the university's policies, regulations, and rules. https://policies.ncsu.edu/

Course Schedule

NOTE: The course schedule is subject to change.

INTRODUCTION (2 LECTURES)

- What are multivariate/longitudinal data? What to expect from this course.
- Graphical displays of data in R
- Basic linear algebra
- Common Distributions

PART I: MULTIVARIATE DATA ANALYSIS (10 LECTURES)

- Principal component analysis
- Factor analysis
- Inference
  - Multivariate normal density. Likelihood function. Maximum likelihood estimation.
  - Inference about a single mean vector (Hoteling's T2, Likelihood ratio test, simultaneous confidence intervals)
  - Inference about two or more mean vectors (paired comparison, repeated measured design, MANOVA, profile analysis)
  - Inference beyond the mean vectors. Tests for the covariance matrix [covariance pattern/totally un-structured]. Tests for multivariate normality [time permitted]
- Classification
- Clustering

PART II: LONGITUDINAL DATA ANALYSIS (10 LECTURES)

- General Linear model
  - Parametric mean model: polynomial, spline, smooth functions where penalties are used to shrink the parameters]
  - Parametric models for covariance
  - Estimation and inference by maximum likelihood
  - Restricted maximum likelihood estimation
- Linear mixed model (LMM)
  - Growth curves
  - Random coefficient model.
  - Estimation, inference, and prediction in linear mixed effects models
- Generalized linear model (GLM)
  - Marginal models, generalized estimating equations (GEE): Estimation and inference of regression parameters
  - Generalized linear mixed models (GLMM): Estimation and inference

ST 537: Applied Multivariate and Longitudinal Data Analysis

NC STATE UNIVERSITY

NORTH CAROLINA STATE UNIVERSITY

RALEIGH, NC 27695

919.515.2011