ST 516 Course Syllabus

ST 516 – Experimental Statistics for Engineers II

Section 001

Fall 2021

3 Credit Hours

Course Description

General statistical concepts and techniques useful to research in engineering. Multiple regression, variable selection, regularization, tree-based regression, experimental design fundamentals, factorial experiments, blocking, random effects, 2k factorials, fractional factorials, response surface methods.

Learning Outcomes

Upon completion of this course students will be able to:

1. Use multiple regression techniques to build interpretable models to predict continuous response variables from continuous and categorical predictor variables.

- 2. Understand assumptions behind linear regression, and condition data to meet those assumptions when necessary.
- 3. Use cross-validation to estimate and minimize model prediction errors.
- 4. Generate predictive models using regularization techniques.
- 5. Implement predictive models using tree-based techniques, including Random Forest and gradient boosting.
- 6. Use R programming language for fitting a variety of predictive models.
- 7. Understand the fundamentals of factorial experimental design and analysis, including blocking techniques and design optimality.

8. Use fractional factorial designs for screening, understanding aliasing behavior and efficient strategies for identifying important factors.

9. Implement response surface methods to identify operating points of important factors that optimize a specific response or group of responses.

10. Use JMP software classical and custom platforms to design and analyze experiments.

Course Structure

Lectures: 2 per week, TuTh 8:30am-9:45am ET with recordings avaiable via Panopto

Homework: 6 total assignments (roughly bi-weekly), solutions submitted electronically via Moodle

Projects: one midterm and one final project, working in tems of 4-5, reports submitted electronically via Moodle

Course Policies

There will be no late work accepted as a general policy. One homework assignment will be dropped to help mitigate unforeseen circumstances that may impact your ability to turn in an assignment on time. In certain instances an extension may be granted if the instructor is contacted at least 24 hours prior to the assignment due date and time.

Instructors

Dan Harris (doharris) - Instructor Email: doharris@ncsu.edu Phone: (828) 446-3635 Office Location: SAS Hall 5118 Office Hours: My office hours are offered in-person in SAS 5118 or may be accessed via Zoom M 1:00pm-3:00pm ET Th 10:30am-12:30pm ET Yuqi Su (ysu25) - Teaching Assistant

Email: <u>ysu25@ncsu.edu</u> Phone: N/A Office Location: Yuqi will offer office hours via Zoom Office Hours: Tu 11:30am-12:30pm ET W 4:30pm-5:30pm ET

Course Meetings

Lecture

Days: TH Time: 8:30am - 9:45am Campus: Main Location: Riddick 451 This meeting is required.

Meeting Notes

All lectures will be recorded and archived on Panopto for students to review. Please be advised this course is being recorded for current and potential future educational purposes. By your continued participation in this recorded course, you are providing your permission to be recorded. Please be aware that the situation regarding COVID-19 is frequently changing, and the delivery mode of this course may need to change accordingly. Regardless of the delivery method, we will strive to provide a high-quality learning experience.

Course Materials

Textbooks

Introduction to Statistical Learning with Applications in R - James Gareth, et. al.

Edition: 1st ISBN: 9781305684164 Web Link: <u>https://catalog.lib.ncsu.edu/catalog/NCSU2798710</u> Cost: Free eBook download from NCSU Libraries *This textbook is required.*

Design and Analysis of Experiments: A Approach - Bradley Jones and Douglas Montgomery

Edition: 1st ISBN: 9781119746010 Cost: \$50 rent, \$120 buy eBook; \$150 buy paperback This textbook is optional.

Expenses

None.

Materials

None.

Requisites and Restrictions

Prerequisites

ST 515

Co-requisites

None.

Restrictions

Graduate Students inside the Engineering Program

Transportation

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

Safety & Risk Assumptions

None.

Grading

Grade Components

Component	Weight	Details
Homework	200	There will be 6 homework assignments worth 40 points each. They will be posted to the course website and will be due the following week by submitting a single pdf or image file with solutions via Moodle in addition to a second file containing R code when appropriate. One lowest homework grade will be dropped. For each assignment, some additional problems will be assigned for extra practice; solutions for these problems to not need to be submitted.

Component	Weight	Details
Midterm Project	200	Midterm Project: posted 9/9/2021, due 10/7/2021. You will work in teams of 4-5. Instructor will provide a data set with one response variable and many potential predictor variables and ask your team to fit a model that has high estimated prediction accuracy and is also intepretable; your team will also be asked to find an optimal operating point given the important predictors you identify. Results will be provided in a written report with a problem statement, description of the modeling process, and conclusions. Requests for re-grading of exams must be made in writing. These requests should contain a complete description of the reason for grade adjustment and the student's name. The request should be attached to the exam and submitted to the instructor within two weeks of the day graded exams are made available to the students on Moodle.
Final Project	200	Final Project: posted 11/4/2021, due 12/2/2021. You will work in teams of 4-5. Instructor will provide problem statement with a response variable for a process and several potentially important factors. Your team is asked to conduct an series of experiements to identify the important factors as well as the operating points the optimize the response variable. Your team will design a series of experiments sequentially and submit the design to the instructor, who will then return simulated data for the response variable of each experimental run. After your team has completed the sequence of experiments, results will be provided in a written report with a problem statement, description of the experimental process, and conclusions. Requests for regrading of exams must be made in writing. These requests should contain a complete description of the reason for grade adjustment and the student's name. The request should be attached to the exam and submitted to the instructor within two weeks of the day graded exams are made available to the students on Moodle.
Options Related to COVID-19	N/A	NC State returned to normal class grading beginning Summer 2021. The "enhanced S/U grading" and "late drop" options are no longer available. For more information, visit <u>https://studentservices.ncsu.edu/your-resources/covid-19/spring2020-sat-grading/#return</u> .

Letter Grades

This Course uses Standard NCSU Letter Grading:

 $97 \leq A+ \leq 100$ 93 ≤ **A** < 97 $90 \leq A - < 93$ 87 ≤ **B+** < 90 83 ≤ **B** < 87 80 ≤ **B-** < 83 77 ≤ **C+** < 80 73 ≤ **C** < 77 70 ≤ **C-** < 73 67 ≤ **D+** < 70 63 ≤ **D** < 67 60 ≤ **D-** < 63 0 ≤ **F** < 60

Requirements for Credit-Only (S/U) Grading

Performance in research, seminar and independent study types of courses (6xx and 8xx) is evaluated as either "S" (Satisfactory) or "U" (Unsatisfactory), and these grades are not used in computing the grade point average. For credit only courses (S/U) the requirements necessary to obtain the grade of "S" must be clearly outlined.

Requirements for Auditors (AU)

Information about and requirements for auditing a course can be found at <u>http://policies.ncsu.edu/regulation/reg-02-20-04</u>.

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 3.17.G at http://www.ncsu.edu/grad/handbook/index.php

Late Assignments

NCSU Course Syllabus: ST 516 - 001 - Experimental Statistics for Engineers II

All due dates are firm and no late work will be accepted, unless arrangements are made with the instructor prior to the time assignments are due. Computer failures, lost files, and sickness (other than COVID-19) or other difficulties are not generally valid excuses for submitting an assignment late. Exceptions to this policy may be made in the event a student is quarantined because of COVID-19; please inform the instructor as soon as possible with appropriate medical documentation to make arrangements for any necessary alterations to assignment schedules.

Attendance Policy

For complete attendance and excused absence policies, please see http://policies.ncsu.edu/regulation/reg-02-20-03

Attendance Policy

Attendance is not generally required, but students are responsible for all information communicated during lectures. Students are strongly encouraged to attend live lecture sessions, and when that is not possible, to review lecture recordings posted on Moodle.

If you test positive for COVID-19, or are told by a healthcare provider that you are presumed positive for the virus, please work with the instructor on health accommodations and follow other university guidelines, including self-reporting (<u>Coronavirus Self Reporting</u>): Self-reporting is not only to help provide support to you, but also to assist in contact tracing for containing the spread of the virus. If you are in quarantine, have been notified that you may have been exposed to COVID-19, or have a personal or family situation related to COVID-19 that prevents you from attending this course synchronously, please connect with the instructor to discuss the situation and make alternative plans, as necessary.

Absences Policy

Since attendance is not generally required, there is no excused absence policy.

Makeup Work Policy

There is no make up for homework assignments unless arranged in advance (see Late Assignments Policy). Students who are unable to attend an exam for a legitimate unavoidable reason may take a make-up exam only if they provide suitable documentation. According to university policy, a student must notify the instructor in advance if s/he will miss an exam. If it is not possible to notify the instructor in advance, the instructor must be given notice as soon as possible after the exam. Suitable documentation of an absence: examples include a physician's note in case of illness or letter from the University or a student's advisor. Students who have a personal emergency (extreme family illness or death, etc.) should contact the Division of Academic & Student Affairs (515-2446; http://dasa.ncsu.edu/) to obtain documentation.

Additional Excuses Policy

If you are quarantined or otherwise need to miss class because you have been advised that you may have been exposed to COVID-19, you should not be penalized regarding attendance or class participation. However, you will be expected to develop a plan to keep up with your coursework during any such absences. If you become ill with COVID-19, you should follow the steps outlined in the Attendance Policy section above. COVID 19-related absences will be considered excused; documentation need only involve communication with your instructor.

Academic Integrity

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

Basic calculators (such as TI-XX) or computer software (such as R) may be used in conjunction with any resources posted to Moodle (lecture recordings, lecture slides, previous assignments, practice exams) or the Internet to complete all homework assignments and projects. Students may form groups to collaborate on homework, but every student must submit their own solutions. For projects, no personal communication with anyone outside your team other than the instructor is permitted regarding the project. Violations of academic integrity will be handled in accordance with the Student Discipline Procedures (NCSU REG 11.35.02).

Academic Honesty

See http://policies.ncsu.edu/policy/pol-11-35-01 for a detailed explanation of academic honesty. None.

Honor Pledge

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

Digital Course Components

Students may be required to disclose personally identifiable information to other students in the course, via digital tools, such as 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.

Digital Course Components:

Moodle: you will need a computer and reliable Internet access.

NCSU Course Syllabus: ST 516 - 001 - Experimental Statistics for Engineers II

Panopto: you will need a computer and reliable Internet access with adequate bandwidth for video streaming. Zoom (office hours only): you will need a computer and reliable Internet access with adequate bandwidth for video streaming, plus web camera, headphones, and microphone.

If you need access to additional technological support, please see the NC State Libraries Technology Lending program at https://www.lib.ncsu.edu/devices.

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 Resource Office at Holmes Hall, Suite 304, 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 (REG02.20.01) (<u>https://policies.ncsu.edu/regulation/reg-02-20-01/</u>).

Non-Discrimination Policy

NC State provides equal opportunity and affirmative action efforts, and prohibits all forms of unlawful discrimination, harassment, and retaliation ("Prohibited Conduct") that are based upon a person's race, color, religion, sex (including pregnancy), national origin, age (40 or older), disability, gender identity, genetic information, sexual orientation, or veteran status (individually and collectively, "Protected Status"). Additional information as to each Protected Status is included in NCSU REG 04.25.02 (Discrimination, Harassment and Retaliation Complaint Procedure). NC State's policies and regulations covering discrimination, harassment, and retaliation may be accessed at https://oied.ncsu.edu/divweb/. 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.

Course Schedule

NOTE: The course schedule is subject to change.

Lecture TH 8:30am - 9:45am - Week 1 - 08/15/2021 - 08/21/2021

ISLR chs. 2.1-2.3: intro to statistical learning, R ISLR ch. 3.1: review of simple linear regression

Lecture TH 8:30am - 9:45am - Week 2 - 08/22/2021 - 08/28/2021

ISLR chs. 3.2-3.3: multiple linear regression ISLR ch. 3.3: predictor interactions, diagnostics

Lecture TH 8:30am - 9:45am - Week 3 - 08/29/2021 - 09/04/2021

ISLR chs. 3.3-3.4: diagnostics, variable transformations ISLR ch. 6.1: subset selection

Lecture TH 8:30am - 9:45am - Week 4 - 09/05/2021 - 09/11/2021

ISLR ch. 5.1: cross-validation ISLR ch. 6.2: regularization (ridge regresssion and lasso) Midterm Project: (assigned 9/9, due 10/7)

Lecture TH 8:30am - 9:45am - Week 5 - 09/12/2021 - 09/18/2021

ISLR ch. 6.3: principal components ISLR ch. 8.1: regression trees

Lecture TH 8:30am - 9:45am - Week 6 - 09/19/2021 - 09/25/2021

ISLR ch. 8.2: bagging, Random Forests, gradient boosting

Lecture TH 8:30am - 9:45am - Week 7 - 09/26/2021 - 10/02/2021

ISLR 4.1-4.3: classification, logisitic regession, classification trees

Lecture TH 8:30am - 9:45am - Week 8 - 10/03/2021 - 10/09/2021

Fall Break Intro to design of experiments, JMP Midterm Project (due 10/7)

Lecture TH 8:30am - 9:45am - Week 9 - 10/10/2021 - 10/16/2021

Simple comparative experiments Single categorical variable experiments

Lecture TH 8:30am - 9:45am - Week 10 - 10/17/2021 - 10/23/2021

Single categorical variable experiments Two-factor factorial experiments

Lecture TH 8:30am - 9:45am - Week 11 - 10/24/2021 - 10/30/2021

Two-factor factorial experiments 2k factorial designs

Lecture TH 8:30am - 9:45am - Week 12 - 10/31/2021 - 11/06/2021

2k factorial designs Blocking in 2k factorial designs Final Project (assigned 11/4, 12/2)

Lecture TH 8:30am - 9:45am - Week 13 - 11/07/2021 - 11/13/2021

Screening designs

Lecture TH 8:30am - 9:45am - Week 14 - 11/14/2021 - 11/20/2021

Response surface designs

Lecture TH 8:30am - 9:45am - Week 15 - 11/21/2021 - 11/27/2021

Covariates

Thanksgiving

Lecture TH 8:30am - 9:45am - Week 16 - 11/28/2021 - 12/04/2021

Final Project (due 12/2)

© 2021 North Carolina State University