DSC 382 Foundations of Regression & Predictive Modeling

Spring 2023 Syllabus

Instructor: Professor Stephen Walker *(He/Him/His)*

Teaching Assistant (TA): Shuying Wang *(She/Her/Hers)*

Learning Facilitators (LFs): Mustapha Saad; Heather Gooss

Office Hours: The majority of communication in this course will be hosted through the Piazza discussion board. TA and LFs will hold virtual office hours via Zoom on a weekly schedule TBA. If you have a grading question or other specific concern that should not be shared with the entire class please email the professor and TAs using the course e-mail address to request an appointment. Along with appointment request, explain the type of query and the professor will determine how best to proceed; whether using an individual appointment, something the TAs will deal with or whether more appropriate for it to be discussed on Piazza.

Course Overview

The course introduces students to the concepts behind the relationships between variables, more commonly known as regression modeling. To develop ideas, the course will begin with the simple linear regression model with single predictor variables and explore the connection with correlation. Developments and extensions are to the multiple linear regression model, combining variable selection and Bayesian approaches, ANOVA models, Generalized Linear models, nonlinear and nonparametric models, regression and classification trees, and mixed effects models. The emphasis throughout will be on estimation algorithms and procedures, hypothesis testing, model selection, goodness of fit, and prediction. Applications will be demonstrated using the R package.

Prerequisites

A background in mathematical statistics and probability is important. An understanding of likelihood methods, and statistical concepts such as hypothesis testing, goodness of fit, and properties of estimators will be essential. Exposure to the package R and Bayesian methods will be useful.

Course Materials

Slides with supporting files including technical derivations omitted from slides. R code for estimation algorithms.

Technical Requirements

Before beginning this course, learners should check that their equipment meets the following technical requirements.

Laptop of personal computer with the following requirements:

- Browser: latest Chrome or Firefox (Chrome preferred).
- OS: Mac 10.12/10.13/10.14, Linux, or Windows 10 with at least 2GB of free space.
- Hardware: Dual Core 2.4 Ghz, 4GB RAM or better.
- Internet connection: cable modem/DSL or better (500 kbps download, 300 kbps upload).

Smartphone or scanner to take pictures and make PDFs of homework submissions.

Smartphone or other device capable of being used for dual-factor authentication.

Homework/Assignments

The course will contain 12 homework/assignments; one per week.

- Each homework/assignment has two problems each in five parts, and includes data analysis.
- Homework/assignments will be a combination of multiple choice (auto grading by edX) and open response (peer marking).
- There will be 5 multiple choice and 7 open response homework/assignments.
- The lowest two homework/assignments will be dropped.

Grading Policy

Percentage to letter grade

A	A-	B+	B	B-	C+	C	C-	D+	D	D-	F
>=90	>=86	>=82	>=78	>=74	>=71	>=68	>=66	>=64	>=62	>=60	<60

Late Policy

No late homework/assignment submissions are accepted.

Program Grade Requirements

30 hour program
9 required hours
21 elective hours

Required courses, B- or higher
Elective courses, C or higher

To graduate, all students must have a graduate GPA avg of at least 3.00.

Course Schedule:

All due dates are 00:01 AM UTC (start of the indicated day, in coordinated universal time). For a conversion to your local time zone see, for example, https://www.timeanddate.com/worldclock/timezone/utc.

Changes to the schedule may be made at the instructor's discretion and if circumstances require. It is your responsibility to note these changes when announced (although we will do our best to ensure that you receive the changes with as much advanced notice as possible).

WEEK 1 (01/07/23)

Foundations and Ideas; Simple Linear Model; Correlation; Estimation; Testing.
HW#1 due 01/14/23

WEEK 2 (01/14/23)

Multiple Linear Regression; Vector and matrix notation; Colinearity; Ridge regression.
HW#2 due 01/21/23

WEEK 3 (01/21/23)

Moderation & Interaction; Testing for interaction; Sobel test; Spotlight on testing.
HW#3 due 02/04/23

REVISION BREAK

WEEK 4 (02/04/23)

Bayes Linear Model; Conjugate model; Prior to posterior analysis; Bayes factor.
HW#4 due 02/11/23

WEEK 5 (02/11/23)

Variable Selection; LASSO; Regularization; Principal components; Bayesian methods.
HW#5 due 02/25/23

REVISION BREAK

WEEK 6 (02/25/23)

ANOVA Models; One-way ANOVA; Two-way ANOVA; ANOVA Table; F-tests.
HW#6 due 03/04/23

WEEK 7 (03/04/23)

Nonlinear Regression; Iterative estimation algorithms; Bootstrap.
HW#7 due 03/18/23

SPRING BREAK

WEEK 8 (03/18/23)

Poisson regression; Analysis of count data; Weighted linear model.
HW#8 due 03/25/23

WEEK 9 (03/25/23)

Generalized Linear Model; Exponential family; GLM theory; Logistic regression.
HW#9 due 04/01/23

WEEK 10 (04/01/23)

Nonparametric Regression; Kernel smoothing; Splines; Regression trees.
HW#10 due 04/15/23

REVISION BREAK

WEEK 11 (04/15/23)

Multiclass Regression; Classification trees; Multinomial logistic regression; Quantile regression; Cox regression.
HW#11 due 4/22/23

WEEK 12 (04/22/23)

Mixed Effects Model; Fixed and random effects; EM algorithm; Gibbs sampler.
HW #12 due 04/29/23

Miscellanea:

Academic Integrity in This Online Course

Each student in the course is expected to abide by The University of Texas Honor Code: "As a student of The University of Texas at Austin, I shall abide by the core values of the University and uphold academic integrity." **Plagiarism is taken very seriously at UT.** Therefore, if you use words or ideas that are not your own (or that you have used in previous class), you must cite your sources. Otherwise you will be guilty of plagiarism and subject to academic disciplinary action, including failure of the course. You are responsible for understanding UT's Academic Honesty and the University Honor Code which can be found at the following web address:
https://deanofstudents.utexas.edu/conduct/standardsofconduct.php

The online course format allows for multiple methods of identity verification, collusion, collaboration and plagiarism monitoring and detection. A violation of the course policy may include (but is not limited to) the following:

- Providing your UT EID to any other person
- Collaborating or sharing information with another person regarding the material on any quiz, assessment or assignment, before, during and/or after any quiz, assessment or assignment
- Recording any quiz, assessment or assignment material in any format
- Failing to properly cite language, ideas, data, or arguments that are not originally yours
- The public (such that it can be viewed by more than one person) posting of any form of a test bank or group of questions from any assignment
- Consulting forbidden materials or sources of information

The University of Texas at Austin Academic Integrity principles call for students to avoid engaging in any form of academic dishonesty on behalf of yourself or another student. Grade-related penalties are routinely assessed ("F" in the course is not uncommon), but students can also be suspended or even permanently expelled from the University for scholastic dishonesty.

If you have any questions about what constitutes academic dishonesty, please refer to the Dean of Students website or contact the instructor for this course.

You must agree to abide by the Honor Code of the University of Texas. You will not work with or collaborate with others in any way while completing any of the graded course assignments.

Sharing of Course Materials is Prohibited:

No materials used in this class, including, but not limited to, lecture hand-outs, videos, assessments (quizzes, exams, papers, projects, homework assignments), in-class materials, review sheets, and additional problem sets, may be shared online or with anyone outside of the class unless you have explicit, written permission. Unauthorized sharing of materials promotes cheating. It is a violation of the University's Student Honor Code and an act of academic dishonesty. We are well aware of the sites used for sharing materials, and any materials found online that are associated with you, or any suspected unauthorized sharing of materials, will be reported to Student Conduct and Academic Integrity in the Office of the Dean of Students. These reports can result in sanctions, including failure in the course. Allegations of Scholastic Dishonesty will be dealt with according to the procedures outlined in Appendix C, Chapter 11, of the General Information Bulletin, http://www.utexas.edu/student/registrar /catalogs/.

Student Rights & Responsibilities

- You have a right to a learning environment that supports mental and physical wellness.
- You have a right to respect.
- You have a right to be assessed and graded fairly.
- You have a right to freedom of opinion and expression.
- You have a right to privacy and confidentiality.
- You have a right to meaningful and equal participation, to self-organize groups to improve your learning environment.
- You have a right to learn in an environment that is welcoming to all people. No student shall be isolated, excluded or diminished in any way.

With these rights come responsibilities:

- You are responsible for taking care of yourself, managing your time, and communicating with the teaching team and with others if things start to feel out of control or overwhelming.
- You are responsible for acting in a way that is worthy of respect and always respectful of others.
- Your experience with this course is directly related to the quality of the energy that you bring to it, and your energy shapes the quality of your peers' experiences.
- You are responsible for creating an inclusive environment and for speaking up when someone is excluded.
- You are responsible for holding yourself accountable to these standards, holding each other to these standards, and holding the teaching team accountable as well.

Personal Pronouns

Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, culture, religion, politics, sexual orientation, gender, gender variance, and nationalities. Class rosters are provided to the instructor with the student's legal name, unless they have added a "preferred name" with the Gender and Sexuality Center, which you can do so here: http://diversity.utexas.edu/genderandsexuality/publications-and-resources/. I will gladly honor your request to address you by a name that is different from what appears on the official roster, and by the gender pronouns you use (she/he/they/ze, etc). Please advise me of any changes early in the semester so that I may make appropriate updates to my records. For instructions on how to add your pronouns to Canvas, visit https://utexas.instructure.com/courses/633028/pages/profile-pronouns.

Documented Disability Statement

The University of Texas at Austin guarantees that students with disabilities have access to appropriate accommodations. You may request an accommodation letter from the Division of Diversity and Community Engagement, Services for Students with Disabilities.

Any student with a documented disability who requires academic accommodations should contact Services for Students with Disabilities (SSD) at (512) 471-6259 (voice) or 1-866-329-3986 (videophone). Faculty are not required to provide accommodations without an official accommodation letter from SSD. Please notify us as quickly as possible if the material being presented in class is not accessible (e.g., instructional videos need captioning, course packets are not readable for proper alternative text conversion, etc.). Reference the SSD website for more disability-related information: https://diversity.utexas.edu/disability/.

Course Etiquette
We expect that you will treat online discussions as though you are having a civil, respectful discussion with your fellow classmates in the same classroom. Please refrain from using profanity or any euphemisms for profanity. Please do not bait other commenters or personally attack them. Please do not use sarcasm in a way that can be misinterpreted negatively. And please do not make the same point over and over again. In short, please just respect the right of your colleagues to ask questions and discuss their opinions about the subject matter of our course on the discussion board. Violators of these discussion rules will simply be shut out from all class communications—email, Piazza, and office hours.

Questions about Technical Problems:
If you have a technical problem, please contact edX support by emailing support-masters@edx.org. Please provide your name, EID, and make sure you identify yourself as a student in this class.

Discussion Forum Language
For technical issues, please contact edX support by emailing support-masters@edx.org. In your email, please provide your name, EID, and the name of the class in which you are having issues.

Additional Language/Guidance
For any technical issues, please contact edX support directly. This may anything related to your edX support or if you have any issues with the edX platform. If you have any questions about your course, including needing extensions, please contact your Teaching Assistant.

Academic Advisor Support
If you have additional questions or require support from an academic advisor, please contact the program coordinator at MSDSGradCoordinator@utexas.edu.

Course Email: onlineregression@austin.utexas.edu