San José State University
Department of Mathematics and Statistics
Math 261A Regression Theory & Methods, Fall 2020

Course and Contact Information

Instructor:	Dr. Guangliang Chen
Email:	guangliang.chen@sjsu.edu
Class Days/Time:	TR 9-10:15am
Meeting Mode:	Online, synchronous (Zoom ID: 952 4846 6497)
Registration Link:	https://sjsu.zoom.us/meeting/register/tJEvc-GsqD0uE9Mh1DI0d8Ngi9hVWwszI7zP
Office Hours:	TWR 10:30-11:30pm (Zoom ID: 983 3362 3851), and by appointment
Piazza:	https://piazza.com/class/kdt29tgm77h2wa
Webpage:	https://www.sjsu.edu/faculty/guangliang.chen/Math261a.html
Prerequisites:	Math 39 and Math 161A (each with a grade of B or better), Math 163* and Math 167R* (*may be taken concurrently)

Course Description

Simple linear regression, multiple regression, indicator variables, influence diagnostics, transformations, assumption analysis, generalized linear models, nonlinear regression, CART, hypothesis testing, confidence and prediction intervals, variable selection and model building. 3 units.

Student Learning Outcomes

After successfully completing this course students should be able to

- Fit simple and multiple linear regression models to data sets.
- Evaluate the adequacy of regression models via residual analysis.
- Compute by hand the parameters of (simple) regression models and/or interpret the output of statistical software to find regression parameters.
- Phrase and test meaningful hypotheses on regression model parameters.
- Evaluate regression models with regards to necessity of transformation of either predictor(s) or response and apply such transformations.
- Identify leverage and influence points in simple and multiple regression applications.
- Fit polynomial regression models in one or two variables and find set of predictor values that optimizes the response.
- Fit regression models to data sets that include one or more categorical predictors. Create indicator variables to code for the categorical predictors. Interpret software output for different methods of contrast coding in multiple regression. Phrase hypotheses for a specific application and create contrasts with which these hypotheses can be tested.
- Understand the relationship between regression and ANOVA models.
- Conduct variable selection by exhaustive search of the space of all predictor subset models. Compare appropriate criteria to evaluate the subset models.
- Apply variable selection algorithms using software and interpret the results.
- Apply the methods learned in the course to the analysis of a data set and present the results of the analysis both in an oral presentation and in the form of a summary paper.

Textbook

Montgomery, Peck & Vining, Introduction to Linear Regression Analysis. 5th edition, 2012. ISBN: 978-0-470-54281-1. This text is available as an e-book (for free) through the SJSU library.

Technology and Equipment Requirements

A scientific calculator is required by homework and exams. Calculators that can compute Normal, t, F, and chi-squared distribution probabilities and quantiles (e.g., TI-84) are preferred. Access to a computer that runs R is required for completing the homework assignments.

Students are required to have an electronic device (laptop or desktop) with a camera and built-in microphone in order to attend the Zoom-based classes and the Proctorio-based test. Additionally, they should have access to a scanner (physical or cell phone app) in order to scan and submit their work.

Students are responsible for ensuring that they have access to reliable Wi-Fi during tests. If students are unable to have reliable Wi-Fi, they must inform the instructor, as soon as possible or at the latest one week before the test date to determine an alternative. See the Learn Anywhere website at https://www.sjsu.edu/learnanywhere/equipment/index.php for current Wi-Fi options on campus.

Any student that needs accommodations or assistive technology due to a disability should work with the Accessible Education Center (AEC), and notify the instructor at the earliest possible time.

Zoom Classroom Etiquette

Please arrive at each Zoom meeting on time. If you have to miss a class occasionally, please find out from the instructor or your classmates regarding what's said/done in that class and act accordingly.

Students are encouraged to turn on their cameras (when without privacy concerns) during each Zoom meeting. If using a virtual background, it should be appropriate and professional and should NOT suggest or include content that is objectively offensive or demeaning.

Please raise your hand to get the instructor's permission before you speak up in class. Alternatively, you can type your question or answer (when responding to the instructor's question) in the chat window.

To help keep background noise to a minimum, make sure you mute your microphone when not speaking.

Lecture Recording

All lectures will be recorded and shared with the whole class; however, you should still make every effort to attend all classes. The recordings will be deleted at the end of the semester. If you would prefer to remain anonymous during these recordings, then please speak with the instructor about possible alternatives.

Students are not allowed to record without instructor permission: Students are prohibited from recording class activities (including class lectures, office hours, advising sessions, etc.), distributing class recordings, or posting class recordings. Materials created by the instructor for the course (syllabi, lectures and lecture notes, presentations, etc.) are copyrighted by the instructor. The university policy (S12-7) is in place to protect the privacy of students in the course, as well as to maintain academic integrity through reducing the instances of cheating. Students who record, distribute, or post these materials will be referred to the Student Conduct and Ethical Development office. Unauthorized recording may violate university and state law.

Course Requirements and Assignments

Course requirements include weekly homework assignments, two projects, and two midterm exams.

Homework will be assigned regularly in Canvas. You can write your work on paper or a tablet (in the former case you need to be able to scan your work). Once completed, submit a legible, electronic copy of your work to Canvas (as a single file attachment).

The course requires two projects, each completed either on your own or along with a classmate (thus to form a group of size two). The first project is to learn a new regression topic and teach it to the rest of the class, while the second project is to apply regression methods learned in this course to a large data set. More details are given in class.

The course has two midterms, both of which are open book, open notes, but you are not allowed to communicate with other people during the test or use the internet to search for answers. More details are given in class.

You are expected to attend all Zoom meetings and actively participate in online discussions. Additionally, you are expected to spend at least 6 hours per week outside of class time on this course.

Grading Information

You must show all necessary steps in your work to earn full credit. Some homework questions require coding in R, in which case you need to provide the R scripts you used, present your results in an organized, meaningful way, and interpret them carefully. Note that it is your work (in terms of correctness, completeness, and clarity), not just your answer, that is graded (correct answers without supporting steps or codes will receive very little credit).

Students may collaborate on homework but must write independent solutions according to their own understanding and styles. Copying and other forms of cheating will not be tolerated and may result in a failing grade for the course, combined with an appropriate disciplinary action from the university.

You must submit homework on time to receive full credit. Late submissions within 24 hours of the due time can still be accepted but will receive a penalty of 10% of the total grade. Submissions that are late for more than one day (24 hours) will not be accepted for any reason. However, your lowest homework score will be dropped.

No make-up exam will be given if you miss a midterm exam unless you have a legitimate excuse such as illness or other personal emergencies and can provide documented evidence.

The two projects will each require a short presentation in class, and Project 2 (the application-oriented one) additionally requires writing a report to present all the technical details. Both of the presentations and report will be graded in terms of depth, clarity, accuracy, and completeness.

The weights in determining the semester total are:

- Homework: 20%
- Midterm 1: 25%
- Midterm 2: 25%
- Project 1: 10%
- Project 2: 20%

Course grades will be assigned by combining the following cutoffs and the actual class distribution (I reserve the right to slightly adjust these percentages in the end):

Grade	*Percentage*	*Grade*	*Percentage*	*Grade*	*Percentage*	*Grade*	*Percentage*
A plus	*97 to 100%*	*B plus*	*86 to 89%*	*C plus*	*73 to 75%*		
A	*93 to 96%*	*B*	*80 to 85%*	*C*	*68 to 72%*	*D*	*60 to 64%*
A minus	*90 to 92%*	*B minus*	*76 to 79%*	*C minus*	*65 to 67%*	*F*	*0 to 59%*

Academic Dishonesty

Students who are suspected of copying in homework or cheating during an exam will be referred to the Student Conduct and Ethical Development office and depending on the severity of the conduct, will receive a zero on the assignment or a grade of F in the course.

University Policies

Per University Policy S16-9 *(http://www.sjsu.edu/senate/docs/S16-9.pdf)*, relevant university policy concerning all courses, such as student responsibilities, academic integrity, accommodations, dropping and adding, consent for recording of class, etc. and available student services (e.g. learning assistance, counseling, and other resources) are listed on Syllabus Information web page (http://www.sjsu.edu/gup/syllabusinfo), which is hosted by the Office of Undergraduate Education. Make sure to visit this page to review and be aware of these university policies and resources.

Math 261A Regression Theory and Methods, Fall 2020, Course Schedule

This schedule is subject to change with fair notice in class and also through Canvas messaging system.

Class	Date		Topic	Textbook
1	AUG 20	R	Introduction and overview	Chapter 1
2	25	T	Simple linear regression	Chapter 2
3	27	R	Simple linear regression	
4	SEP 1	T	Simple linear regression	
5	3	R	Multiple linear regression	Chapter 3
6	8	T	Multiple linear regression	
7	10	R	Multiple linear regression	
8	15	T	Model adequacy checking	Chapter 4
9	17	R	Model adequacy checking	
10	22	T	Model adequacy checking	
11	24	R	Review & Project 1 presentations	
12	29	T	Midterm 1	
13	OCT 1	R	Transformations and weighting	Chapter 5
14	6	T	Transformations and weighting	
15	8	R	Diagnostics for leverage and influence	Chapter 6
16	13	T	Polynomial regression models	Chapter 7
17	15	R	Polynomial regression models	
18	20	T	Polynomial regression models	
19	22	R	Indicator variables	Chapter 8
20	27	T	Indicator variables	
21	29	R	Indicator variables	
22	NOV 3	T	Multicollinearity	Chapter 9
23	5	R	Variable selection and model building	Chapter 10
24	10	T	Variable selection and model building	
25	12	R	Review & Project 1 presentations	
26	17	T	Midterm 2	
27	19	R	Validation of regression models	Chapter 11
28	24	T	Generalized linear models	Chapter 13
29	DEC 1	T	Some special topics (if time permits)	Chapter 15
30	3	R	Project presentations	
Final exam day	9	W	Project 2 presentations (7:15-9:30am)	