Stat 435[M]: Statistical Modeling for Data Analytics

Fall 2022

3 Credits (2 Lectures, 1 Lab)

Instructor:	Xiongzhi Chen
Office:	Neill Hall Room 230
Contact means:	Office phone at (509) 335-3150, and email at xiongzhi.chen@wsu.edu
Lecture times and location:	14:10 – 13 on Mondays and Wednesdays at EME B0046
Lab time and location:	15:10 – 17 on Fridays at CUE 114
Office hours:	1pm – 1:50pm on Mondays and Wednesdays, 1pm – 2:20pm on Fridays, or by appointments
Prerequisites:	STAT 360
Required material:	***An Introduction to Statistical Learning with Applications in R*** (2017), G. James, D. Witten, T. Hastie, R. Tibshirani. Available from Amazon.com or in free pdf online at *https://statlearning.com/*
Grader:	Madison Neyland

Description

The course will cover the following topics and their software implementations via R:

- Linear regression models in low-dimensional settings, their inferences, and their diagnostics.
- Linear regression models in high-dimensional settings, model selection, and post-selection inference. This includes classical model selection methods and two shrinkage methods (the LASSO and ridge regression), and a very brief introduction to post-selection inference techniques for LASSO and ridge regression.
- Resampling techniques ("cross validation" and "bootstrap") and their applications in linear regression and other settings.
- Logistic regression and classification techniques via discriminant analysis.
- Non-linear regression such as polynomial regression, smoothing splines, and generalized additive models.

Student Learning Outcomes

At the end of the course the student will	**Schedule**
Be able to apply linear regression models, make inference on model parameters, assess model assumptions, and address violations of model assumptions.	Week 3-6
Be able to use bootstrap and cross-validation.	Week 5-9
Be able to conduct model selection such as subset selection, stepwise selection, and the LASSO, and apply two shrinkage methods: Ridge regression and the LASSO. Be able to conduct post-selection inference for LASSO and Ridge regression.	Weeks 7-9
Be able to apply logistic regression, and conduct classification tasks using discriminant analysis.	Week 10-12
Be able to use nonlinear regression models such as polynomial regression, smoothing splines, and generalized additive models.	Weeks 13-16

Assessment of Learning Outcomes

Students' learning outcomes will be accessed by a combination of homework assignments, quizzes, projects and exams. Detailed information is given by the following.

Grading: Grading will be based on quizzes, homework assignments, written projects, and a final exam (if any). These elements are outlined in the following paragraphs.

Quizzes: Approximately 3 quizzes will be given during the semester. The quiz problems will be developed from materials presented in lectures and the textbook. These will primarily consist of programming exercises. All quizzes are conducted in an open book and open notes manner. The points from all quizzes will comprise 25% of the final grade.

Make-up Quizzes: Make-up quizzes will be allowed on a case-by-case basis and will be given to accommodate university conflicts, illness or other unforeseen emergencies. Make-up quizzes are not guaranteed, and there must be a strong reason for the need to make up a quiz. Students must let the instructor know, as soon as possible, that they will not be able to take the scheduled quiz. Make-up quizzes must be completed before the WSU official final exam date(s) for the semester of the course and within a reasonable period after they were originally scheduled.

Homework: Approximately 5 homework assignments will be given during the semester. These will come from problems provided by the textbook or materials discussed in the lectures. Homework assignments will primarily consist of methodological exercises and programming exercises, and both the two parts of exercises need to be completed. Please submit answers to HW assignments with necessary supporting computer codes, and organize them. The points from all homework assignments will comprise 40% of the final grade. Late homework will only be accepted under extenuating circumstances, such as an extended illness.

Written Projects: One project and one final exam (or a second project) will be assigned during the semester. The project(s) and final exam will comprise 35% of the final grade. Each project will consist of 4 components: (1) a typed write-up that contains 5 components: introduction, methods used to conduct the analysis, results of the analysis, conclusions from and discussion on the analysis, and a reference section; (2) computer codes used to conduct the analysis; (3) most relevant outputs from the analysis (which can be incorporated in component (1)); (4) a detailed description of the contribution of each member of the group towards the project. **Each project can be completed by up to 2 students.**

Written Project Grading Rubric:

Introduction: This section must consist of a description of the problem to be addressed, a concise review of pertinent literature related to the problem, and a statement of the proposed solution to the problem. It should usually not exceed 2 pages in length.

Methods: This section should describe how the problem described in *Introduction* is solved. It usually is the main part of the project report.

Results: This section should present the main results of the solution to the problem and their interpretations.

Discussion: This section should discuss aspects of the solution that can be further improved, relevant issues not touched by the solution, and other problems to which the solution might be applied.

Class Participation/Attendance: In order to properly gain a working knowledge of the material, attendance and participation in class are necessary. While this will not be graded, enrolled students are highly encouraged to attend the class.

Grade Distribution:

Homework	40%
Quizzes	25%
Written Projects/ Final exam	35%
Total	100%

Evaluation Criteria: The grading scale for this course will be as follows:

A	93% - 100%	C+	77% - 79.99%
A-	90% - 92.99%	C	73% - 76.99%
B+	87% - 89.99%	C-	70% - 72.99%
B	83% - 86.99%	D+	66% - 69.99%
B-	80% - 82.99%	D	60% - 65.99%
		F	0% - 59.99%

Other important information

Students with Disabilities: Reasonable accommodations are available for students with documented disabilities or chronic medical or psychological conditions. If you have a disability and need accommodations to fully participate in this class, please visit your campus' Access Center/Services website to follow published procedures to request accommodations. Students may also contact their campus offices to schedule an appointment with a Disability Specialist or an Access Advisor. All disability related accommodations are to be approved through the Access Center/Services on your campus. It is a university expectation that students visit with instructors (via email, Zoom, or in person) to discuss logistics within two weeks after they have officially requested their accommodations. For more information contact a Disability Specialist on your home campus:

(a) Pullman, WSU Global Campus, Everett, Bremerton, and Puyallup: 509-335-3417 **Access Center** (https://www.accesscenter.wsu.edu) or email at **access.center@wsu.edu**
(b) Spokane: 509-358-7816 **Access Services** (https://spokane.wsu.edu/studentaffairs/access-resources/) or email **j.schneider@wsu.edu**
(c) Tri-Cities: **Access Services** (http://www.tricity.wsu.edu/disability/) or email **g.hormel@wsu.edu**
(d) Vancouver: 360-546-9238 **Access Center** (https://studentaffairs.vancouver.wsu.edu/student-wellness-center/access-center) or email **van.access.center@wsu.edu**

Academic Integrity: Academic integrity is the cornerstone of higher education. As such, all members of the university community share responsibility for maintaining and promoting the principles of integrity in all activities, including academic integrity and honest scholarship. Academic integrity will be strongly enforced in this course. Students who violate WSU's Academic Integrity Policy (identified in Washington Administrative Code WAC 504-26-404) will receive a score of zero on any graded coursework, which may result in failing the course, will not have the option to withdraw from the course pending an appeal, and will be reported to the Center for Community Standards.

Cheating includes, but is not limited to, plagiarism and unauthorized collaboration as defined in the Standards of Conduct for Students, WAC 504-26-010(3). You need to read and understand all of the definitions of cheating: **http://app.leg.wa.gov/WAC/default.aspx?cite=504-26-010**. If you have any questions about what is and is not allowed in this course, you should ask course instructors before proceeding.

If you wish to appeal a faculty member's decision relating to academic integrity, please use the form available at **communitystandards.wsu.edu**. Make sure you submit your appeal within 21 calendar days of the faculty member's decision.

Safety Statement: Classroom and campus safety are of paramount importance at Washington State University, and are the shared responsibility of the entire campus population. WSU urges students to follow the "Alert, Assess, Act," protocol for all types of emergencies and the **"Run, Hide, Fight"** response for an active shooter incident. Remain ALERT (through direct observation or emergency notification), ASSESS your specific situation, and ACT in the most appropriate way to assure your own safety (and the safety of others if you are able).

Please sign up for emergency alerts on your account at MyWSU. For more information on this subject, campus safety, and related topics, please view the **FBI's Run, Hide, Fight video** and visit the **WSU safety portal**. Full details can be found at **https://provost.wsu.edu/classroom-safety/**

Policy Prohibiting Discrimination and Harassment (Executive Policy 15)**:** This policy expresses WSU's commitment to maintaining an environment free from discrimination, including sexual harassment. This policy applies to all students, faculty, staff, or others having an association with the University. Discrimination, including discriminatory harassment, sexual harassment, and sexual misconduct (including stalking, intimate partner violence, and sexual violence) is prohibited at WSU (See WSU Policy Prohibiting Discrimination and Harassment (Executive Policy 15) and WSU Standards of Conduct for Students).

If you feel you have experienced or have witnessed discriminatory conduct, you can contact the WSU Office of Compliance and Civil Rights (CCR) and/or the WSU Title IX Coordinator at 509-335-8288 to discuss resources, including confidential resources, and reporting options. (Visit https://ccr.wsu.edu/ for more information).

Most WSU employees, including faculty, who have information regarding sexual harassment or sexual misconduct are required to report the information to CCR or a designated Title IX Coordinator or Liaison. Visit https://ccr.wsu.edu/file-acomplaint/ for more information.

Expectations for Classroom Conduct: Respect each other and treat others how you want to be treated. Please silence your cell phones and all other electronics and refrain from using these items during class. Do not disrupt the class, students are here to learn and cannot do so if others are being disruptive. If I feel you are disrupting the class or are disrespectful of anyone, I reserve the right to ask you to leave class for the day. Success in class requires reading the textbook, listening and asking questions in lectures, and doing all assigned work. Only you choose whether or not to succeed by doing these things.

Class Attendance and Absences (Academic Regulation 72)

Per Academic Regulation 72, students are responsible for ensuring that they attend all class meetings and complete all in-class and out-of-class work as assigned by the instructor. Students are also responsible for communicating with the instructor should they need to be absent.

- Administrative Drops for Non-Attendance

  Students who have not attended class meetings (including lectures, laboratories, and other meetings) during the first week of the semester or according to a prorated schedule for shorter sessions may be dropped from the course by the department. Students enrolled in online classes may be dropped if they have not logged into the class during the first week.

  Students should not assume that they have been dropped without checking their class schedules.

- Absences

Students should make all reasonable efforts to attend all class meetings. However, in the event a student is unable to attend a class, it is the responsibility of the student to inform the instructor as soon as possible, explain the reason for the absence (and provide documentation, if appropriate), and make up class work missed within a reasonable amount of time, if allowed. Missing class meetings may result in a reduction in a student's overall grade in the class.

- University Sponsored. Any student who is required to participate in off-campus, university-sponsored activities such as field trips, musical performances, judging teams, intercollegiate athletic events, etc., should obtain an official Class Absence Request form from the faculty or staff member supervising the on- or off-campus activity. The form must contain specific information concerning the activity and date, be signed by the supervising faculty or staff member, and be submitted by the student at least one week in advance to the individual instructors of the student's classes. Alternative forms of the class Absence Request may also be provided by the supervising faculty or staff member and are acceptable for requesting university sponsored absences. It is recommended, but not required, that a student not be penalized for absence from class provided a properly signed form has been filed with the instructor prior to the absence. These university-sponsored absences are subject to an instructor's attendance policy and are not intended to imply additional acceptable absences. In all instances, it is the student's responsibility to make up all work missed.

- Military Service Members. Students who are members of the National Guard or a reserve branch of a military service are occasionally required to miss class for weekend drills, active duty, and related responsibilities and are entitled to make up the class (RCW 28B.10.270). In such a case, instructors must not penalize students for the absences and must allow them to make-up the missed work. In each instance, it is the responsibility of the student to inform the instructor of the duty before the absence, provide appropriate documentation if requested, and complete the missed work as soon as reasonably possible.

- Reasonable Religious Accommodation. Washington State University reasonably accommodates absences allowing for students to take holidays for reasons of faith or conscience or organized activities conducted under the auspices of a religious denomination, church, or religious organization. Reasonable accommodation requires the student to coordinate with the instructor on scheduling examinations or other activities necessary for course completion. Students requesting accommodation must provide written notification within the first two weeks of the beginning of the course and include specific dates for absences. Approved accommodations for absences will not adversely impact student grades. Absence from classes or examinations for religious reasons does not relieve students from responsibility for any part of the course work required during the period of absence. Students who feel they have been treated unfairly in terms of this accommodation may refer to Academic Regulation 104 - Academic Complaint Procedures. See also Academic Regulation 82, available at https://registrar.wsu.edu/academic-regulations/

- Adverse Weather Conditions. When appropriate campus authorities declare a "Yellow / Delayed or Limited Operations" or "Red / Closed" campus condition (BPPM 50.40-46), or travel to instructional locations is unsafe, classes may be cancelled or delayed. When a student does not attend due to adverse conditions, the instructor will not penalize the student. See the Inclement Weather Policy online at each campus website and at the Office of the Provost website.

- Other Absences. Students must sometimes miss class meetings, examinations, or other academic obligations affecting their grades due to extenuating circumstances. It is the

responsibility of the student to provide a written explanation for the absence to the instructor as soon as it is reasonable to do so. When possible, students should provide appropriate documentation for their absence but instructors cannot require written excuses from health care professionals.

As long as absences are not excessive, it is recommended, but not required, that the instructor provide and document reasonable arrangements. Determinations regarding the acceptance of an absence are the discretion of the instructor based on the attendance policy as stated in the class syllabus. Students who attempt to gain advantage through abuse of any aspect of the absence policy (e.g., by providing an instructor with false information) may be referred to the Center for Community Standards.

Policy on Recordings: Without the written content from the Instructor, a student should not modify or distribute recordings of lectures, lab practices or office hours, and should not use any such recording for any commercial purpose.

Complaints: Students who wish to raise a concern about the instructor's arrangement regarding missed work may follow the Academic Complaint Procedure, outlined in Academic Regulation 104.

Tentative Schedule of Events

Week 1
Lectures - Course overview. Review of probability and statistics (Instructor supplied material).
Lab - Basics of R: basic objects and operations; import and export data; basic commands to process data.

Week 2
Lectures - Continuation of the review of probability and statistics (Instructor supplied material).
Lab - Basics of R: generate random variables, and obtain basic statistical quantities; basic plot commands.

Week 3 Homework 1 assigned.
Lectures - Simple linear regression (including qualitative predictors), its inference, and model diagnostics (Chapter 3 and Instructor supplied materials).
Lab - Using R to conduct simple linear regression, its inference, and model diagnostics.

Week 4
Lectures - Multiple linear regression (including qualitative predictors), its inference, and model diagnostics (Chapter 3 and Instructor supplied materials).
Lab - Using R to implement multiple linear regression, its inference, and model diagnostics. Project 1 assigned.

Week 5 (Note: Lecture on Wednesday becomes Lab practice.) Homework 2 assigned and Homework 1 due.
Lecture 1- Model diagnostics for linear regression (Chapter 3 and Instructor supplied materials)
Lecture 2- Using R to implement model diagnostics for linear regression.
Lab - Quiz 1.

Week 6
Lectures - The bootstrap procedure and its applications (Chapter 5 and Instructor supplied materials).
Lab - Using R to implement bootstrap and its application to linear regression.

Week 7 Homework 3 assigned and Homework 2 due.
Lectures - Model selection methods for linear regression models: subset selection, stepwise selection, and the LASSO. (Chapter 6).
Lab - Using R to implement subset selection and stepwise selection.

Week 8
Lectures - Shrinkage methods: LASSO and Ridge regression. (Chapter 6).
Lab - Using R to implement LASSO and ridge regression.

Week 9 (Note: Lecture on Wednesday becomes Lab practice)
Lecture 1 - Ridge regression, cross-validation and post-selection inference (Chapter 5, Chapter 6, and Instructor supplied materials).
Lecture 2 - Using R to implementation cross-validation and post-selection inference for LASSO and ridge regression.
Lab - Quiz 2.

Week 10 Homework 4 assigned and Homework 3 due. (Discussion of whether to have another project or have an open book, open notes final exam.)
Lectures - Simple logistic regression, its inference, and model diagnostics. (Chapter 4 and Instructor supplied materials).
Lab - Using R to implement simple logistic regression, its inference, and model diagnostics.

Week 11 Project 1 due, and Project 2 assigned (if required).
Lectures - Multiple logistic regression model, its inference, and model diagnostics. (Chapter 4 and Instructor supplied materials).
Lab - Using R to implement multiple logistic regression, its inference, and model diagnostics.

Week 12
Lectures - Classification via discriminant analysis: Gaussian mixture models, and discriminant analysis. (Chapter 4).
Lab - Using R to implement discriminant analysis.

Week 13 Homework 5 assigned and Homework 4 due.
Lectures - Non-linear regression methods: polynomial regression and smoothing splines. (Chapter 7)
Lab - Using R to implement polynomial regression and smoothing splines.

Week 14 **Thanksgiving break**

Week 15 (Note: Lecture 2 on Wednesday becomes a Lab practice.)
Lecture 1 - Non-linear regression methods: smoothing splines and generalized additive models. (Chapter 7)
Lecture 1 - Using R to implement smoothing splines and generalized additive models.
Lab - Quiz 3.

Week 16 Final Exam (if required); Project 2 due (if required). Homework 5 due.

Disclaimer: The schedule and procedures outlined in this syllabus are subject to change depending on circumstances.