Advanced Regression

Fall 2025 Block 3

Instructor

Instructor Dr. Tyler George

Cornell College, West 311

tgeorge@cornellcollege.edu

Class Meetings

October 20th to November 12th

Daily, 9am-11am and 1pm-3pm

West 201

Office Hours

MWTh 3:05pm-4:05pm and by appt.

West 311

Optional Appointment

You Are A Priority

My goal for this block is to help you learn the material. I want to first and foremost recognize that you are an individual and thus unique, and may learn uniquely. Additionally, your health and well-being are priority one. Learning cannot happen effectively if you don't meet your other personal needs. That being said, I have structured the class in a way that, based on my experience teaching and learning, I believe will be most beneficial for most students. I promise to do my best to create an inclusive and engaging learning environment. I kindly request that you maintain an open line of communication with me when you need help or require flexibility. You and your learning are why I am here.

Course Description

Following a second regression course, this class will begin with a review of multiple linear regression, but now using R. New topics will include probability distributions, likelihoods, differentiating binary vs binomial logistic regression, and poisson regression including its variants. The class of generalized linear models will then be presented which unifies all past modeling approaches. All methods are presented using realistic case studies. Conducting and communicating the modeling process including exploratory data analysis, model exploration and selection, and inferences are all emphasized. Prerequisite(s): STA 202 and DSC 223. Alternate years.

Learning Objectives

This course supports the Educational Priorities and Outcomes of Cornell College with emphasis on knowledge, communication, and reasoning. Specifically, the learning objectives of this course are: - Learn statistical distributions, likelihood functions, and types of models, including poisson regression, quasi-logistic regression, multilevel models, and polynomial regression. - Ability to communicate statistical ideas clearly and accurately (communication). - Understand the class of generalized linear models and the models contained (knowledge, reasoning). - Ability to ascertain which type of analysis is appropriate for a particular data set and question (reasoning).

Prerequisite

To be successful in this class, you should have completed STA 201, STA 202, and DSC 223.

Open Access Books – Free!

All of the materials for this class are free.

The main textbook is: Beyond Multiple Linear Regression by Paul Roback and Julie Legler – it is freely available online. Chapters 1-9.

The secondary text is: An Introduction to Statistical Learning with Applications in R by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani – it is freely available online. Chapter 7.

Course Site and Moodle

Our course will run from a combination of Moodle and the course website at https://stats-tgeorge.github.io/STA363_AdvRegF25/.

Software – No need to install

We will use a combination of technologies in this course, including R and RStudio (server). I have set up an RStudio server on a machine located on campus that we will all access using a web browser. You don't need to install any software (in fact, I prefer you not to, for a while)—more on this in class. If you are an off-campus student, please notify me immediately, as there are additional steps required for setting up access to your work off campus.

You can access the RStudio Server at: <http://turing.cornellcollege.edu:8787/>.

If you encounter any technical issues, please contact IT as soon as possible. Submit a Work Order!

```
### Group Work

#In this class, I would like you to work in groups for a variety of reasons. A large part of
```

Evaluations and grades

Grade Category Descriptions

Homework:

Homework will be graded for correctness. I will generally evaluate a random subset of the assigned questions. The goal is to practice applying the method and then be able to interpret the results.

Labs and Class Participation

I will measure this by your class attendance and your work on labs and class examples. With class work, please save your script files in the folder that I can view on the RStudio server. A folder will be created for you on the server (details will be provided in class). Generally, labs are credit/no credit, but with some feedback provided in class. I will review these files to give participation points at the end of the block.

Mini-Projects

This class includes a variety of related techniques; generally, only one of them is the correct technique for a particular research question. For this reason, we will do labs following a single concept and mini "projects" at the end of each primary new method or collection of methods. With labs, the data will be provided. With mini-projects, you may need to find and clean data. The mini projects will be primarily completed in class and will include exploratory analysis, identifying, fitting, and evaluating models, interpreting results, a short presentation, and a brief report. Completed in groups.

Exams

You will have two exams this block, Due on October 31st and November 12th. Each will have two components. Component 1 will be held on these dates. Component 2 will be a take-home, open-book, open-note exam, distributed beforehand, and due on the specified dates. You will have 12 hours or more to complete the take-home portion.

Assignment	Points
Homework	300
Labs & Class Participation	75
Mini Projects	225
Exams, two 200pts exams	400
Total	1000

Grade	Range	Grade	Range
A	93-100%	C	73-76%
A-	90–92%	C-	70-72%
B+	87–89%	D+	67-69%
B	83-86%	D	63-66%
B-	80-82%	D-	60-62%
C+	77-79%	F	<60%

AI Policy

Work, including writing, code, and analysis (such as graphs, tables, models, etc.), created by AI tools will not be considered original work and will instead be considered automated plagiarism. It is derived from previously created texts from other sources that the models were trained on; however, it doesn't cite these sources.

AI models have built-in biases (ie, they are trained on limited underlying sources; they reproduce, rather than challenge, errors in the sources) AI tools have limitations (ie, they lack critical thinking to evaluate and reflect on criteria; they lack deductive reasoning to make judgments with incomplete information at hand; they make up or use inaccurate information and may "hallucinate" sources that do not exist)

In this course, the use of AI is only permitted when specified. You can use it as a learning tool, but no submissions should contain any text or code written by a Large Language Model (e.g., ChatGPT, Gemini). These tools are not familiar with the content of this class, and they will provide incorrect answers. By taking this class, you are learning content that it does not know. Writing tools, such as Grammarly, are allowed. If you are unsure if the tool you want to use is OK, please ask.

DISABILITIES AND ACCOMODATIONS POLICY

Cornell College makes reasonable accommodations for persons with disabilities. Students should notify the Office of Academic Support and Advising and their course instructor of any disability related accommodations within the first three days of the term for which the accommodations are required, due to the fast pace of the block format. For more information on the documentation required to establish the need for accommodations and the process of requesting the accommodations.

ACADEMIC HONESTY POLICY

Cornell College expects all members of the Cornell community to act with academic integrity. An important aspect of academic integrity is respecting the work of others. A student is expected to explicitly acknowledge ideas, claims, observations, or data of others, unless generally known. When a piece of work is submitted for credit, a student is asserting that the submission is her or his work unless there is a citation of a specific source. If there is no appropriate acknowledgment of sources, whether intended or not, this may constitute a violation of the College's requirement for honesty in academic work and may be treated as a case of academic dishonesty. The procedures regarding how the College deals with cases of academic dishonesty appear in The Catalog, under the heading "Academic Honesty."

Illness Policy

If you are experiencing COVID-19 symptoms, do not attend class. Perform a home test or contact Director of Student Health Services Lynn O'Brien at **student_health@cornellcollege.edu** immediately to arrange a COVID-19 test at the Health Center.

Mandatory Reporter Reminder

It is my goal that you feel supported and able to share information related to your life experiences during classroom discussions, in your written work, and in any one-on-one meetings with me. You should also know that all Cornell College faculty and staff are mandatory reporters. This means that I will keep information you share with me private to the greatest extent possible. However, I am required to share information regarding sexual assault, abuse, criminal behavior, or about a student who may be a danger to themselves or to others. If you wish to speak to someone confidentially who is not a mandatory reporter, you can schedule an appointment with one of the counselors in the Ebersole Health and Wellbeing Center or contact the College Chaplain, Rev. Melea White, at **mwhite@cornelllcollege.edu**.