STOR 455H: Methods of Data Analysis (Honors): Spring 2026

This course teaches data analysis methods with a primary focus on linear regression. Linear regression is the starting point for all more advanced statistical tools that aim to predict one "response" variable based on a number of other variables known as covariates. It is very widely used in medical and public health applications, social sciences including economics, and applications in biological and environmental sciences. More advanced techniques such as generalized linear models, survival analysis, longitudinal data analysis, etc., all require a basic foundation in linear regression. Therefore, a sound understanding of linear regression is essential for anyone seeking to use advanced statistical methods in almost any context.

Compared with the non-Honors version of this course, STOR 455, the aim of this course is to go into much more detail about the mathematical foundations of the subject, but still with focus on computational applications in R as in the current STOR 455. Mathematical prerequisites include the basic undergraduate calculus sequence and an introductory course in linear algebra such as Math 347.

Outline of topics covered:

- Introduction to linear regression. Simple linear model, principle of least squares, derivation of least squares estimates. Multiple regression, reformulation in matrix terms and derivation of least squares estimates by matrix methods.
- Means, variances, covariances and distributions of estimators. Confidence intervals, hypothesis tests, prediction intervals.
- Diagnostics: Residuals, including standardized and studentized residuals. Outliers and influential observations. Measures of influence including DFFITS, DFBETAS and Cook's Distance.
- Variable selection. Best subset, forward, backward and stepwise regression. Model criteria including adjusted R-squared, Mallows Cp, AIC, BIC. Transformations. Multicollinearity.
- ANOVA models: one-way, two-way, factorial designs, randomized blocks.
- Implementation and practical applications in R.

Texts:

1. *Linear Models with R*, Third Edition (2025), Julian J. Faraway, Chapman and Hall/CRC Press.
2. *Linear Regression* by Richard L. Smith. Instructor's draft text, in preparation for publication, will be available on the course canvas page. PLEASE UDERSTAND THAT THIS IS A PRELIMINARY VERSION AND IS NOT TO BE CIRCULATED OUTSIDE THE CLASS.

Instructor:

Richard L. Smith, Department of Statistical and Operations Research. Office: Hanes 303. Email: rls "at" email "dot" unc "dot" edu. Website: https://rls.sites.oasis.unc.edu/.

Office hours (subject to change): Mondays 1:00-3:00 pm; Wednesdays 12:30-1:30 pm; Thursdays 1:30-2:30 pm. I am willing to meet by zoom or at other times by arrangement – feel free to email me to request an appointment.

In case we do have zoom meetings, the link will be https://unc.zoom.us/j/8238989389.

Instructional Assistant and Grader:

Anna Myakushina (amyak "at" unc "dot" edu). Office hours: Tuesday 11:00-12:00 in Hanes B30

Course time and place:

Tuesdays and Thursdays 9:30-10:45, Hanes 107.

The first class is on Thursday, January 8, and the last class is on Thursday, April 23. There will be no class on Tuesday, March 17 or Thursday, March 19 (Spring Break), or Thursday, April 2 (Well Being Day). If you need accommodation on other holidays that I am not aware of, please discuss it with me.

Prerequisites:

The official prerequisite is to have completed either STOR 120 or STOR 155. However, it is very much to be hoped that most of you will have taken at least an introductory statistics course (AP statistics would be acceptable) plus the basic calculus sequence and some exposure to linear algebra (MATH 347 would certainly qualify for that). In the first few classes I will try to assess your degree of preparedness in these topics and will, if necessary, make adjustments to the teaching materials.

Class Attendance Policy

This is an in-person class and attendance at all class sessions is expected of all students.

If you are not able to attend class for any reason, please email me in advance (or in emergency, as soon as you are able to after the class). Acceptable excuses include illness, medical appointments, job interviews, major family events such as weddings, and religious holidays. In particular, please do not attend class if you have COVID or some other infectious disease. There is no need to obtain an official university excuse unless you expect to miss at least three classes.

We will have quizzes! You must be present in class to take the quiz and to receive credit for it. At my discretion, I *may* allow you to take the quiz remotely if you email me promptly with an acceptable excuse (see previous paragraph), but it's your responsibility to ask for this. Otherwise you will simply receive a score of zero for the missed quiz.

Quizzes are not intended to be difficult or time consuming. They will mostly be based on material that was covered in the previous class.

Face Masks and Vaccinations

There is no requirement that students at UNC must either wear a face mask or be vaccinated. If you have an infection or are concerned about picking one up, you are welcome to wear a face mask.

Assignments and Exams

The grading of the course will be split among quizzes (15%), homeworks (20%), two midterm exams (20% each) and final exam (25%). Homeworks will be given at weekly or bi-weekly intervals and will cover the mathematical content of the course as well as numerical exercises to be completed in R or RStudio. You may use R-Markdown (if you are familiar with it) for your assignments but this is not required; alternative formats are latex or Word (e.g. copy your figures and tables into a Word document and add explanations as appropriate). Mathematical assignments may be completed by hand or in latex or Word (your choice), but if possible, please scan and convert to pdf. Assignments will be announced on the course canvas page and are to be handed in via gradescope through the gradescope link on canvas.

As noted above, quizzes will be short reviews of recent material expected to be completed in class. It's possible there won't literally be a quiz every class, but you should always expect that the class will start with a quiz.

The lowest three quizzes and the lowest one homework will be dropped when computing your final grade.

When submitting homeworks using gradescope, we ask you to:

1. Assign pages for each question (if you submit the whole assignment in a single file) or submit separate solutions for each question or part-question;

2. Highlight your final statement for each question or sub question (e.g. use the yellow highlighter feature in R-Markdown or Word) or write your final statement of each question or sub question before all your code and explanations; above all, *always answer the question*;

3. When you have a question about grading, we suggest submitting a regrade request on gradescope stating your concerns instead of sending us email;

4. Check the deadline of each homework to minimize the number of students who miss the deadline.

There will be two midterm exams, both in-class. The tentative dates for these exams are Tuesday February 10 and Thursday March 12; if you are aware of a conflict with either of those dates, please let me know a.s.a.p.

The final exam is scheduled by the Registrar's office for 8:00 – 11:00 am on Thursday, May 7. Unless announced otherwise, please assume this will be an in-class exam held in Hanes 130. Any changes to these arrangements will be discussed in advance with the class.

You are reminded that the university Honor Code is in effect for this course. Quizzes are expected to be completed in class and to be your own individual work – you may consult the course text or your own notes but not outside resources such as google or any AI platform. For homework assignments, you are allowed to discuss the problems among yourselves, but the work you hand in must be your own; direct copying is not permitted. For exams, whether in-class or take-home, you are expected to work the problems entirely by yourselves and consultation of any kind is forbidden, unless it is with me or the Instructional Assistant.

Assignment of Grades

Unless announced otherwise, the assignment of grades will be as follows:

93-100 A; 90-92.99 A-; 87-89.99 B+; 83-86.99 B; 80-82.99 B-; 77-79.99 C+; 73-76.99 C; 70-72.99 C-; 60-69.99 D+; 50-59.99 D; 0-49.99 F.

At the instructor's discretion, these grade point boundaries may be lowered, but they won't be raised.

Policy on AI

As you are all aware, AI tools such as ChatGPT or Microsoft Copilot have become very powerful and there is much discussion of their appropriate place in the educational setting. Different professors have established different policies and other professors in the STOR department, and elsewhere in the university, may have conditions very different from mine, but the following points will apply in this course:

- Homework assignments: AI use is permitted subject to certain conditions. You are allowed to use AI to look up the details of a statistical procedure, background information about a mathematical or scientific topic, or to get general advice about how to tackle problems of the sort discussed in the course. Really, anything you might have looked up in a book or with older tools such as the google search engine. However, the work you hand in must be your own: solutions generated entirely by AI are not acceptable, and may result in an honor code violation if we determine that the deception was deliberate. You should also be aware that AI sometimes produces answers that are irrelevant to the questions at hand or just plain wrong. You are responsible for checking your answers and verifying their correctness: any errors are assumed to be yours, not the machine's. You should include a brief written acknowledgement if you use AI for anything other than a total routine task.

- Quizzes and Exams. No use of AI is permitted on a quiz or an exam. If you violate this policy in a quiz you will get an automatic score of zero for that quiz. On an exam, the same policy will apply but you may also be reported to the Honor Court.

Emergency Preparedness

As you are probably all aware already, there have been increasing concerns about campus safety following some incidents that caused the entire campus to be shut down for a few hours. Here are some resources that I recommend you use in case such incidents occur again.

1. Please refer to the student version of the campus safety webpage - https://campussafety.unc.edu/carolina-ready/be-prepared/students/
2. If you do not already have your cellphone registered for "Carolina Alert" texts, please do that: https://alertcarolina.unc.edu/register/
3. Download the "Carolina Ready Safety App", https://campussafety.unc.edu/safetyapp/. I believe you will receive messages through that app in the event of an emergency.

Regarding specific circumstances that may cause me to suspend the class, while it's impossible to anticipate all eventualities, I believe there are three major kinds of disruptions we should be prepared for:

1. Fire alarm. If the alarm goes off, I will stop teaching, please exit the building immediately and do not return until given the all clear. There are three readily accessible exists from Hanes 130 (any of the three doors to outside on the first floor), you can use any of these.
2. "Shelter in place", which could refer to a tornado or some similar event that requires us to get out of harm's way. If that should happen, probably the safest place to gather will be in the basement of Hanes Hall. You should move there, and again, stay there until someone tells you it is safe to move.
3. "Secure in place", which means we need to stay put and protect ourselves against a possible attacker. If that happens, we will lock the door to Hanes 130, lower the window shades, and take whatever protection measures we can until given the all clear.

Accessibility Resources

The University of North Carolina at Chapel Hill facilitates the implementation of reasonable accommodations, including resources and services, for students with disabilities, chronic medical conditions, a temporary disability or pregnancy complications resulting in barriers to fully accessing University courses, programs and activities.

Accommodations are determined through the Office of Accessibility Resources and Service (ARS) for individuals with documented qualifying disabilities in accordance with applicable state and federal laws. See the ARS Website for contact information: https://ars.unc.edu or email ars@unc.edu. (source: https://ars.unc.edu/faculty-staff/syllabus-statement)

Counseling and Psychological Services (CAPS)

CAPS is strongly committed to addressing the mental health needs of a diverse student body through timely access to consultation and connection to clinically appropriate services, whether for short or long-term needs. Go to their website: https://caps.unc.edu/ or visit their facilities on the third floor of the Campus Health Services building for a walk-in evaluation to learn more. (source: Student Safety and Wellness Proposal for EPC, Sep 2018)

Title IX Resources

Any student who is impacted by discrimination, harassment, interpersonal (relationship) violence, sexual violence, sexual exploitation, or stalking is encouraged to seek resources on campus or in the community. Please contact the Director of Title IX Compliance (Adrienne Allison – Adrienne.allison@unc.edu), Report and Response Coordinators in the Equal Opportunity and Compliance Office (reportandresponse@unc.edu), Counseling and Psychological Services (confidential), or the Gender Violence Services Coordinators (gvsc@unc.edu; confidential) to discuss your specific needs. Additional resources are available at safe.unc.edu.

Honor Code: (For the complete honor code, please visit http://instrument.unc.edu/)

It shall be the responsibility of every student enrolled at the University of North Carolina to support the principles of academic integrity and to refrain from all forms of academic dishonesty, including but not limited to, the following:

1. Plagiarism in the form of deliberate or reckless representation of another's words, thoughts, or ideas as one's own without attribution in connection with submission of academic work, whether graded or otherwise.

2. Falsification, fabrication, or misrepresentation of data, other information, or citations in connection with an academic assignment, whether graded or otherwise.

3. Unauthorized assistance or unauthorized collaboration in connection with academic work, whether graded or otherwise.

4. Cheating on examinations or other academic assignments, whether graded or otherwise, including but not limited to the following:

(a) Using unauthorized materials and methods (notes, books, electronic information, telephonic or other forms of electronic communication, or other sources or methods);

(b) Violating or subverting requirements governing administration of examinations or other academic assignments;

(c) Compromising the security of examinations or academic assignments;

(d) Representing another's work as one's own; or

(e) Engaging in other actions that compromise the integrity of the grading or evaluation process.

5. Deliberately furnishing false information to members of the University community in connection with their efforts to prevent, investigate, or enforce University requirements regarding academic dishonesty.

6. Forging, falsifying, or misusing University documents, records, identification cards, computers, or other resources so as to violate requirements regarding academic dishonesty.

7. Violating other University policies that are designed to assure that academic work conforms to requirements relating to academic integrity.

8. Assisting or aiding another to engage in acts of academic dishonesty prohibited in the above items.

Administrative details

- All questions regarding course registration and waiting list should be directed at Ms. Christine Keat, crikeat@email.unc.edu.
- The instructor reserves to right to make changes to the syllabus.