SC321: Statistical Modeling

Professor: Jerzy Wieczorek
Office: Davis 213
Telephone: 859-5841
Email: jwieczko@colby.edu

Office hours: See our Moodle page

Class meetings: MWF 12:00 – 12:50 in Davis 308 (Computer Lab)

Class materials: Available at our Moodle course site


Course description:
Statistics is at the center of advances in science in the 21st century. Almost every area of modern science depends on statistics in some form. It serves not only the sciences, but the field’s influence has grown to affect all aspects of our lives from health to business to public policy to entertainment. Thanks to advances in technology, we now live in a world inundated with data. How can we make use of all this data to yield useful information? Statistical modeling provides a way of making sense of data, a way to gain insight about processes and populations. Extracting useful information requires more than learning a laundry list of methods; it requires careful consideration of context. What is the question of interest and what theory motivates the investigation? How was the study designed? How were the data collected? What can ultimately be learned from the analysis, that is, what is the scope of inference?

In this course, we will explore several flexible and powerful approaches to statistical modeling, including simple linear regression, multiple regression, model diagnostics, analysis of variance (both simple and multifactor), and logistic regression for binary response variables and binomial counts.

Course objectives:

By the end of the semester:

1. Students will be able to extend linear regression concepts learned in an introductory statistics course to a setting where there is more than one predictor variable.

2. Students will learn how to formulate research questions using statistical models.

3. Students will understand the implications of bias, variation, confounding, and interaction in the context of statistical inference.

4. Students will learn how to check regression model assumptions and assess the fit of different types of models.
5. Students will be able to model some types of data that violate the standard linear regression assumptions (e.g., data with binary outcomes).

6. Students will be able to identify proper regression modeling strategies depending on the type of data encountered.

7. Students will be comfortable using software to fit statistical models and use models to answer research questions.

8. Students will demonstrate the ability to communicate statistical results effectively.

**Grades:** Your course grade will be determined as follows (subject to change):

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class participation</td>
<td>2.5%</td>
</tr>
<tr>
<td>Homework</td>
<td>15.0%</td>
</tr>
<tr>
<td>Exam #1</td>
<td>27.5%</td>
</tr>
<tr>
<td>Exam #2</td>
<td>27.5%</td>
</tr>
<tr>
<td>Final Project</td>
<td>27.5%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100.0%</strong></td>
</tr>
</tbody>
</table>

**Moodle:** I’ll be posting all HW assignments on our Moodle site along with hints, schedule updates, and other important class-related info. Check the Moodle site early and often at [https://moodle.colby.edu](https://moodle.colby.edu)

**Homework:** Homework represents an important aspect of the course. Readings and selected problems will be assigned regularly. The assigned problems and due dates will appear in the Course Schedule on Moodle. Homework will require extensive use of the computer package RStudio. The assignments are designed to thoroughly test your knowledge of topics we discuss in class and are definitely not meant to be one-night jobs. Check Moodle for between-class announcements, hints, or modifications in assignments. Although you are encouraged to discuss problems with each other, I expect each person to submit his or her own work. You may choose to work with a partner on data analyses and thus produce the same computer output – that’s okay, but I want written interpretations and responses to be in your own words. And remember, an answer to a statistics question is almost never just a number – I want thoughtful explanations and conclusions!

Homework due dates and times will be posted on Moodle. Turn in hard-copy homework assignments at my office by 4pm on the due date. **Late homework will reduce the student’s grade substantially (by 25% of the maximum score for each day late), and very late homeworks may not be accepted**, in order to let the graders return everyone’s work quickly and give you timely feedback. The professor will not grant extensions except for documented illness or other extreme circumstances.

**Exams:** There will be two exams during the term. These exams will focus on your ability to use the statistical software, to interpret results, to express an understanding of statistical
concepts, and to engage in statistical thinking on open-ended questions. The exams will involve both in-class and take-home components. See our Moodle page for the two exam dates. **You must meet with me within the first two weeks of classes if you need to reschedule an exam date or request accommodations.**

**Data analysis project:** Your final exam will consist of a data analysis project that you will work on with another classmate. Your grade will be based on project assignments throughout the semester.

**Class sessions:** Regular attendance is essential and in-class participation is expected – please take it seriously. Everyone will receive a grade for in-class participation (0, 1, 2, 3, 4, or 5) that will be incorporated into your final grade. You need to attend class and you need to be prepared for class. Check the Moodle site for reading assignments prior to class meetings. At times, class will involve activities based on these readings. This class brings together students with varying backgrounds in science, math, and statistics, so it is especially vital that you raise questions in class when you don’t understand a particular topic, or if you have a perspective on a topic that you feel would contribute to class discussion. Please bring your textbook with you to class; we will occasionally work through problems in small groups, often using R on our classroom computers.

**Classroom computers:** Our class will meet in a computer lab. **These computers are to be used for educational purposes only.** Students who use computers during class time for purposes beyond SC321 (email, internet surfing, paper writing, etc.) are obviously not participating in class, and this will be reflected in your final grade. Your hands should be off of the keyboard unless you’re using R or our Moodle page.

**Software:** We’ll be using the software package RStudio to analyze data. Today, virtually all data analysis is done with the aid of at least one computer software package (phew!) – R is one of the most powerful and most flexible ones – it’s also probably the most widely used of all statistical packages. You can access RStudio via the web at: [https://rstudio.colby.edu](https://rstudio.colby.edu)

R is becoming more common in almost every field of research. Having a working knowledge of this important tool will come in handy if you ever have to make sense of data in the future. I strongly recommend that you use the webserv ed version of R to ensure that we are all working off of the same version. However, if you wish to install R and RStudio on your own machine, you can first download R (the programming language itself) here: [https://cran.r-project.org/](https://cran.r-project.org/)

After installing R, you can download RStudio (a development environment for R) here: [https://www.rstudio.com/](https://www.rstudio.com/)

In addition, for exams you will also need a scientific calculator (one that can at least do square roots, logs, and exponential functions).
Academic accommodations:
I am available to discuss academic accommodations that any student with a documented disability may require. Please note that you’ll need to provide a letter from the Dean of Studies Office documenting your approved accommodations. Please meet with me within two weeks of the start of the semester to make a request for accommodations so that we can work together with the College to make the appropriate arrangements for you. Kate McLaughlin, Associate Director of Access and Disability Services (kmclaug@colby.edu) is the primary contact for accommodations and any questions related to educational testing and documentation.

Sexual misconduct/Title IX statement:
   Colby College prohibits and will not tolerate sexual misconduct or gender-based discrimination of any kind. Colby is legally obligated to investigate sexual misconduct (including, but not limited to sexual assault and sexual harassment).
   If you wish to speak confidentially about an incident of sexual misconduct, please contact Colby Counseling Services (207-859-4490) or the Director of the Gender and Sexual Diversity Program, Emily Schusterbauer (207-859-4093).
   Students should be aware that faculty members are considered responsible employees; as such, if you disclose an incident of sexual misconduct to a faculty member, they have an obligation to report it to Colby’s Title IX Coordinator. “Disclosure” may include communication in-person, via email/phone/text, or through class assignments.
   To learn more about sexual misconduct or report an incident, visit https://www.colby.edu/sexualviolence/

Academic honesty & consequences for academic dishonesty:
   Honesty, integrity, and personal responsibility are cornerstones of a Colby education and provide the foundation for scholarly inquiry, intellectual discourse, and an open and welcoming campus community. These values are articulated in the Colby Affirmation and are central to this course. Students are expected to demonstrate academic honesty in all aspects of this course.
   Academic dishonesty includes, but is not limited to: plagiarism (including quoting sources without " " around the borrowed words and a citation); presenting another’s work as one’s own; buying or attempting to buy papers or projects for a course; fabricating information or citations; knowingly assisting others in acts of academic dishonesty; violating clearly stated rules for taking an exam or completing homework; misrepresentations to faculty within the context of a course; and submitting the same work, including an essay that you wrote, in more than one course without the permission of instructors.
   Academic dishonesty is a serious offense against the college. Sanctions for academic dishonesty are assigned by an academic review board and may include failure on the assignment, failure in the course, or suspension or expulsion from the College.
   For more on recognizing and avoiding plagiarism, see: https://libguides.colby.edu/avoidingplagiarism
   For resources and information on academic integrity, see: https://www.colby.edu/academicintegrity/
The Colby Affirmation:

Colby College is a community dedicated to learning and committed to the growth and well-being of all its members. As a community devoted to intellectual growth, we value academic integrity. We agree to take ownership of our academic work, to submit only work that is our own, to fully acknowledge the research and ideas of others in our work, and to abide by the instructions and regulations governing academic work established by the faculty. As a community built on respect for ourselves, each other, and our physical environment, we recognize the diversity of people who have gathered here and that genuine inclusivity requires active, honest, and compassionate engagement with one another. We agree to respect each other, to honor community expectations, and to comply with College policies. As a member of this community, I pledge to hold myself and others accountable to these values.

https://www.colby.edu/catalogue/2014/07/11/colby-affirmation/