­­Statistical Methods

Fall 2017

**Instructor**: Michael Lopez

**Office**: Harder 217

**Email:** mlopez1@skidmore.edu

**Twitter:** @StatsbyLopez

**Phone**: 580-5297

**Office Hours:** Thursday 10-12 or by appointment

**Class Meetings:** TR 9:10 – 11:00

**Texts:** *Introductory Statistics with Randomization and Simulation* by Diez et. Al (required)

**Link to download:** http://www.openintro.org/stat/textbook.php

**Link to purchase (Amazon):** http://www.amazon.com/dp/1500576697

**Website**: https://github.com/statsbylopez/StatsMethods

*Mathematical Statistics with Resampling and R*  by Chihara and Hesterberg (optional)

*Modern Dive*  by Ismay and Kim (http://moderndive.com/)

**Course Goals**: Statistics is the science of collecting and interpreting data. It’s a branch of mathematics impacting countless other subjects, including psychology, economics, and biology. That in mind, the points of this course include:

(1) The development and understanding of the most common statistical tools, including confidence intervals, hypothesis testing, linear and multiple linear regression, and bootstrapping

(2) To be able to distinguish real world relevance of terms like statistical significance, experiment design, and interpretation of results

(3) Apply statistical tools to answer questions in whatever topics students are interested in (medicine, public health, sports, nutrition, economics, etc)

**Course Skills:**

-Recognition of benefits of, drawbacks to, and how best to analyze experiments and studies

-The role of resampling in determining statistical significance

-Use data types to determine and conduct the proper type of analysis

-Computing skills using *R* & *R-studio*

-Reproducible research using RMarkdown

**Textbook**: The textbook is free to download and available for a nominal fee on Amazon.com. I encourage you to purchase a hard copy as soon as you can. See the links in the syllabus heading. We will also be taking samples of chapters from other textbooks.

**Computing:** The use of the R statistical environment (downloadable from <http://www.r-project.org/>) with the RStudio interface (downloadable from www.rstudio.org) is thoroughly integrated into the course. Access on or off campus can be round at: <http://r.skidmore.edu>

You will need to bring a scientific calculator to each exam. You will not be allowed to share a calculator with another student. You do not need a graphing calculator: any calculator that can take square roots, logs and exponents will suffice.

**Writing:** Your ability to communicate results, which may be technical in nature, to your audience, which is likely to be non-technical, is critical to your success as a data analyst. The assignments in this class will place an emphasis on the clarity of your writing.

**Grading:**

**Homework [15%]** Homework is the most effective way to reinforce concepts learned in class. There will be weekly homework assignments. Most often, questions will relate to material in the reading that will be covered in class. Homework is due at the start of class, and will be accepted with a 33% penalty if turned in within 48 hours (and no credit otherwise). All homework and solutions are posted on Blackboard.

Homeworks are graded out of 5 points:

1-3 out of 5 points: Most questions attempted, minimal effort

4 of 5 points: All questions attempted, complete effort, graded questions incorrect

4.5 of 5 points: All questions attempted, complete effort, graded questions partially correct

5 of 5 points: All questions attempted, graded questions perfect

**Projects & Presentation [25%]** The major milestone in this class will be conducting (with a group) a statistical investigation on a question of interest to you. For each, you may collect primary data by hand or you can use data available on the Internet or from faculty research.

You will prepare a project proposal describing your study and obtain approval from me before you begin the investigation. During the last week of class, you (and your group) will give a 10 minute oral presentation of your study. We will spend time in class looking at what data is available on the web and about writing a project proposal.

**Exams [50%]** There will be two exams. Exams are take homes, with 24 hours given to complete. Each are expected to take three hours, and you are not allowed to seek outside help on the exams.

**Participation [10%]** Active participation in class, engagement with the annotated readings notes, regular attendance, and completion of in-class labs will comprise the remainder of your grade.

**Disability accommodations**: Any student with special needs requiring accommodations should give me his/her memo of accommodations in a timely manner. It is the student’s responsibility to follow up with me regarding all accommodations that require my participation. The student is advised to ensure full use of testing accommodations by coming to talk to me at least three days before any test.

**Attendance**: Your attendance in class is crucial, as is your punctuality. We are all going to learn this material together, so we need to have everyone present and working. Accommodations for an unavoidable absence can be made in advance via email; one necessary absence during the semester is not unusual; having more than two is uncommon.

**Reading**: See the detailed syllabus for the sections you are responsible. The book is nearly free! Use it, write on it, bring it on vacation, and appreciate it.

**Collaboration:** Much of this course will operate on a collaborative basis, and you are expected and encouraged to work together with a partner or in small groups to study, complete homework assignments, and prepare for exams. However, every word that you write must be your own. Copying and pasting sentences, paragraphs, or blocks of *R* code from another student is not acceptable and will receive no credit. No interaction with anyone but the instructor is allowed on any exams or quizzes. All students, staff and faculty are bound by the Skidmore College Honor Code.

**Title IX Statement:** Skidmore College is committed to providing a learning, working, and living environment that reflects and promotes personal integrity, civility, and mutual respect. Members of the Skidmore community have the right to be free from all forms of abuse, assault, harassment, and coercive conduct, including sexual and gender-based misconduct, as defined in this policy. Skidmore College considers sexual and gender-based misconduct to be one of the most serious violations of the values and standards of the College. Unwelcome sexual contact of any form is a violation of students’ personal integrity and their right to a safe environment and therefore violates Skidmore’s values. Skidmore College will not tolerate sexual or gender-based misconduct in any form. Sexual and gender-based misconduct is also prohibited by federal regulations. In accordance with Title IX, Skidmore College does not discriminate on the basis of sex in any of its programs and activities.

**Title IX Reporting Responsibilities:** Skidmore College faculty are committed to supporting our students and upholding gender equity laws as outlined by Title IX. Therefore, if a student chooses to confide in a member of Skidmore’s faculty or staff regarding an issue of sexual or gender-based misconduct, that faculty or staff member is obligated to tell Skidmore’s Title IX Deputy Coordinator. The Title IX Deputy Coordinator will assist the student in connecting with all possible resources for support and reporting both on and off campus.

**Additional policies:**

Cell phones: If your phone rings, I get to answer it! Respect the rest of the class and turn your phone on silent, airplane mode, or off

Computers: You may bring your own laptop to use for labs, or use the ones provided in Harder. I strongly recommend that computers be closed during note taking, and please be aware that your computer use is a distraction for your classmates. If you would like to use a computer during class, please sit in the back row.

**Calendar (subject to change)**

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Reading | Topics | Assignments Due |
| 9-7 | 1.1-1.2 | Introduction to statistical inference, **Lab 0** |  |
| 9-12 | 1.3-1.4 | Data & experiments, numerical data |  |
| 9-14 | 1.4-1.6 | Categorical Data, **Lab 1** | HW 1 |
| 9-19 | 1.7, 5.1 | Ch. 1 review, linear regression & bivariate data |  |
| 9-21 | \*\* | Multivariate data, **Lab 2** | HW 2 |
| 9-26 | 2.1 | Randomization case studies |  |
| 9-28 | 2.2-2.4 | Simulation, **Lab 3** | HW 3 |
| 10-3 | 2.3-2.4 | Normal distribution |  |
| 10-5 | 2.6-2.7 | Normal Model, **Lab 4** | HW 4 |
| 10-10 | 2.8 | Confidence Intervals |  |
| 10-12 | \*\* | Review for midterm, **Lab 5** | HW 5 |
| 10-17 |  | **Midterm 1 (take home)** |  |
| 10-19 | 3.1 | Inference for 1 proportion | Project Roster |
| 10-24 | 3.2 | Inference for 2 proportions |  |
| 10-26 | 4.1 | 1 sample means, **Lab 6** | HW 6/Proposal |
| 10-31 | 4.2 | Paired data |  |
| 11-2 | 4.3 | Difference of two means, **Lab 6** | HW 7/Prop revision |
| 11-7 |  | Review |  |
| 11-9 | 5.2 | OLS fitting Outliers & regression inference, **Lab 6** | HW 8 |
| 11-14 | 5.3 | Outliers & regression inference |  |
| 11-16 | 6.1 | Multiple regression, **Lab 7** | HW 9/Project data |
| 11-21 | 6.2 | Diagnostics & model selection, project grading |  |
|  |  |  |  |
| 11-28 | 6.3 | Regression diagnostics, **Lab 8** |  |
| 11-30 | \*\* | Review and project work |  |
| 12-5 | \*\* | **Midterm 2 (take home)** |  |
| 12-7 | \*\* | Project work |  |
| 5-2 |  | Project presentations |  |
|  |  | **Technical reports** |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |