MATH 546
Spring 2022
January 10 - May 6
Applied Statistics I
Online

Ranjan Rohatgi
rrohatgi@saintmarys.edu
574-284-4498
Office Hours: virtual/in-person
M9-11, W2-4
Office: MA332

The instructor reserves the right to make changes to the syllabus as deemed necessary. Students will be notified in a timely manner of any syllabus changes via email or on Blackboard.

Course Description: In general, statistics is concerned with the development and application of methods and theories for collecting, analyzing, and interpreting data in such a way that the reliability of the conclusions based on the data may be evaluated objectively by the use of probability theory. As such, it is a relatively new branch of mathematics with much of it being developed in the last 100 years. It is sometimes referred to as *inferential statistics* in order to differentiate it from the common use of the word, which is then called *descriptive statistics*. For example, when we speak of the statistics involved in a baseball game we are using descriptive statistics. Thus, this course is primarily the study of inferential statistics as it is related to probability theory. The course will include theory, but will focus primarily on the application of statistics, specifically as it relates to data science.

Bulletin Description: An introduction to the foundations and applications of statistics. Topics include basic concepts of data collection, sampling, and experimental design; descriptive analysis and graphical displays of data; probability concepts and expectations; normal and binomial distributions; sampling distributions and the Central Limit Theorem; confidence intervals and hypothesis testing; likelihood-based statistics; ANOVA; correlation and simple linear regression.

Program Outcomes: This course satisfies the following MSDS program outcomes:

C. The graduate develops and implements data analysis strategies based on theoretical principles and detailed knowledge of the underlying data.

F. The graduate rigorously applies mathematical principles to the analysis of data.

Course Outcomes: Upon successful completion of the course, a student will be able to:

1. define and use correctly statistical terms, language, nomenclature, and methods introduced in the course;

2. define sampling distribution; know the sampling distributions related to the normal distribution; recognize when a sampling distribution for sample mean and sample proportion are approximately normal, including the use of the Central Limit Theorem;

3. define point estimator and calculate its bias; compute point estimates using the method of maximum likelihood, method of moments, and Bayesian estimation;

4. demonstrate understanding of confidence intervals and hypothesis testing; use these concepts correctly for statistical inference related to population mean, proportion, variance, and independence of variables;

5. perform an analysis of variance (ANOVA) for one- and two-way layouts;

6. analyze categorical data using chi-square test and contingency tables;

7. differentiate between parametric and nonparametric statistical techniques, and understand when to use which type in a given setting;

8. select and use statistical methods required to solve real-world problems;

9. effectively use statistical software to solve statistical problems;

10. communicate ideas and work clearly, correctly, and with precision.

Textbooks: There is no required textbook for this course. Instead, listed below are various optional resources. The instructor will indicate recommended weekly reading assignments from each source listed below to accompany the week's topics.

- *Mathematical Statistics and Data Analysis*, Edition 3, John A. Rice. ISBN: 978-0-534-39942-9

- *Mathematical Statistics with Resampling and R*, Edition 2, Chihara & Hesterberg. ISBN: 978-1-119-41653-1

- *OpenIntro Statistics*, Edition 4, `https://www.openintro.org/book/os/`

Statistical Software: This course will utilize `R`, a free software environment for statistical computing and graphics, and **RStudio**, an integrated development environment for `R`. Both programs are freely available online and the links for downloading the software are provided on Blackboard.

Blackboard: `http://bblearn.saintmarys.edu`
The Blackboard site for this course will provide all necessary course information and documents, including video lectures and assignments. Online quizzes and exams will be administered through the Blackboard site, details below. Discussion Forums will be available on the Blackboard site as an additional way to communicate with the instructor and the class. *Students should access the class Blackboard site frequently, if not daily.*

Getting help: Make an appointment to meet with me! Additionally, work with your classmates on assignments.

Anticipated Weekly Schedule: The following is subject to change at the discretion of the instructor.

Week	Unit	Topics
1		survey sampling: parameters, simple random sampling
2		sampling distribution of $\bar{X}$, CLT, distributions from normal
3		Confidence Intervals, Exploratory Data Analysis
4	1	Parameter Estimation: methods of moments and maximum likelihood
5		Bayesian Estimation
6		**Exam 1**, project EDA due
7		Introduction to Hypothesis Testing
8		Comparing Two Samples: Parametric Tests
9		Comparing Two Samples: Nonparametric Tests
10		One-Way ANOVA
11	2	Two-Way ANOVA
12		**Exam 2**, project test ideas due
13		Analysis of Categorical Data: Chi-Square Tests
14		Analysis of Categorical Data: Exact Test; Matched-Pairs Designs; Odds Ratios
15		work on project
16		**Exam 3**, final project due

Assignments: Students should expect to be working *at least* **12 hours per week** on this course (including reading, study, watching videos, homework, and quizzes). There will be regular reading assignments and problem-solving assignments; students are responsible for remaining current, and for asking questions when they arise. Students are expected to do the assigned reading as scheduled. Problem-solving assignments will be given weekly and posted on Blackboard. Details regarding assignment submission are provided on Blackboard. Since much of the learning process will occur in the process of doing this work, the grade for problem-solving assignments will be included in the final course grade. Students are encouraged to work and study together, but each student is responsible for his or her own work and understanding. Discussion-board assignments and DataCamp assignments will be graded for completion only.

Quizzes: There will be regular quizzes administered online through Blackboard. Each quiz will be based on the material covered in lecture videos, assigned reading, and homework assignments. Students are allowed to use their notes, any of the recommended textbooks, and statistical software to take quizzes. Quizzes are an individual activity though, and students are expected to do all of the work themselves with no help from classmates.

Project: Students will complete a data analysis project and will prepare a report detailing the steps of the project and summarizing the results. Assessment of the project will not only relate to the mathematical and statistical content, but also to the actual write-up, including grammar and formatting.

Testing: There will be three (3) exams given during the semester (note the timing in the Anticipated Weekly Schedule above). Each exam will have two parts: a 25-minute oral exam with me, and an untimed, open-book, open-note, written exam. The final exam is not cumulative.

Grading: Course grades will be based on the following:
    10% Quizzes
    15% Assignments
    15% Project

25% Exam 1
25% Exam 2
10% Exam 3

The grading scale will follow:

$$50\% \leq D < 64\% \leq C\text{-} < 68\% \leq C < 72\% \leq C\text{+} < 76\%$$

$$\leq B\text{-} < 80\% \leq B < 84\% \leq B\text{+} < 88\% \leq A\text{-} < 92\% \leq A$$

Attendance, Participation, and Due Dates: This course is fully online, so we will never meet in a physical classroom. However, "attendance" is still required. Students are expected to **attend** the course by logging into Blackboard, viewing and reviewing provided content (i.e., lecture notes and video lectures), completing assigned work, and to **participate** in the course by communicating with other students in the course via Discussion Forums, and interacting with the instructor. All of this should be done within the Blackboard environment. Students should expect to interact with Blackboard a minimum of four times per week for extended periods of time on each visit.

If you have planned to travel during the course, you should also plan to have access to the course materials (a computer which meets our online course requirements, reliable internet connection, Blackboard) and spend the expected amount of time working on the course as described above.

Blackboard monitors student involvement in the course. Students who visit Blackboard infrequently or without spending the time required to review the necessary materials and interact with the course will be identified and contacted about their attendance in this sense.

All work must be submitted by the posted due date. If a student misses a deadline without prior notification to the instructor, then she/he will receive no credit for the work. Students may request an extension for submitting homework assignments, but may be assessed a late penalty. No extensions will be granted for quizzes or exams.

Online Code of Conduct: The main way in which students will engage in this course and interact with the instructor and classmates is through the discussion forums. In order to establish a healthy and vibrant online learning community all students are expected to adhere to the following discussion forum etiquette:

1. **Be respectful.** Please respect your fellow classmates. Debate and challenging ideas is part of a healthy and thriving learning community, but only when it is done in a polite and respectful manner. Insulting, condescending, or abusive words will not be tolerated and will be removed.

2. **Be constructive.** A learning community is about learning with and through engagement with one another. When commenting on each other's posts, try to offer constructive feedback and/or suggestions for improvement.

3. **Be sensitive.** Your fellow classmates may come from many different cultures and backgrounds. Be mindful of that fact and take care when discussing race, religion, gender, sexual orientation, or controversial topics, since others may be more sensitive about them than you.

The main goal of the Discussion Forums is to provide a place where dialogue can build as the collective learning community digs into the content together. The forums are there for all students

to have a deeper, more meaningful learning experience, but there a few things to consider when crafting a post:

1. ***Read before posting.*** A comment or question should build on what's already been discussed. If you have a question, scan through to see if it's already been addressed and avoid duplicating threads.

2. ***Use informative titles.*** Keep your post titles short and on-topic so that the forums are easy to navigate.

3. ***Make it easy to read.*** Check your posts for spelling or grammatical errors.

4. ***Stay on topic.*** Don't change the topic of a thread or split a thread into multiple discussions. Don't post the same post/question in more than one forum.

5. ***Help keep the site friendly.*** If you think a comment is mean, insulting, or otherwise inappropriate, let the instructor know.

Institutional Syllabus: The Institutional Syllabus can be found at `https://www.saintmarys.edu/academics/resources/registrar`.

Canceled Class: Students will receive an email, sent to their Saint Mary's College email address, notifying them of the cancellation of any synchronous course activities.

Academic Honesty Policy: Students are expected to abide by the Saint Mary's College Academic Integrity and Academic Honesty policy detailed in the *Graduate Bulletin*. Any violation of this policy will be reported to the MSDS Program Director and the student will receive an F for the course. In particular, you may not work with anyone else on quizzes or exams, and any written work you submit must be entirely your own.

Student Support Services: The information about graduate student resources is available at `http://grad.saintmarys.edu/student-resources`.

Accessibility Support: Students eligibile for accommodations should contact Iris Giamo in the Disabilities Resource Office (103C Madeleva Hall, phone 284-4262, email igiamo@saintmarys.edu) for an appointment to review documentation and arrange for appropriate accommodations. Students who suspect they may have a disability are also encouraged to contact the Disabilities Resource Office.

Technology Requirements: Students are required to have access to a working computer with the following minimum hardware specifications.

Operating System:	Windows 7 (Ultimate or Professional) or later or Mac OS X v.10.6 or later
Processor:	One of the following: 1.5GHz Pentium M or better, Intel Core Duo/Core 2 Duo, Intel Core i5/i7 (Windows) or Intel Core Duo/Core 2 Duo, Intel Core i5/i7 processor (MAC)
Memory:	Minimum of 8GB of RAM
Hard Drive:	Minimum of 80GB of storage space

Web Browser: Most recent version of at least one major web browser

Other: Reliable high-speed Internet connection, scanner or other technology (e.g., Livescribe pen) to convert hand written notes to a PDF document, Webcam and microphone, MS Office (covered by Microsoft Campus Agreement) or another text editor (able to convert files to PDF)

File Backup: Students are responsible for keeping a backup copy of course related files on their computer, Google Drive, portable storage devices, etc.. Lost files do not excuse missing or late work.

Note: Students are responsible for having a working computer that satisfies all the minimum requirements

Technology Skills

Students should be able to communicate through discussion boards, and use (or be able to learn quickly) Web conferencing tools and the use of a scanner (or other technology capable of converting their handwritten work into a PDF document). Students should have sufficient technical proficiency to be able to learn new software applications and new technical tools.

Technical Support

For technical assistance, please contact ResNet, the Saint Mary College IT Helpdesk.

Email: resnet@saintmarys.edu
Phone: (574) 284-5319
Office Location: 113 Haggar Campus Center (on Saint Mary's campus)
Website: http://sites.saintmarys.edu/~resnet