Syllabus

Site:	Rose-Hulman Institute of Technology	Printed by:	Megan Heyman
Course:	MA382: Fall 2020-21	Date:	Thursday, February 4, 2021, 3:46 PM
Book:	Syllabus		

Description

** I reserve the right to modify the course content, schedule, topics, policies, etc. outlined in this syllabus.**

Table of contents

Basic Information

This document may be changed or adjusted as needed during the term

Catalog Description

MA382 is an introductory course in statistics. Dual emphasis is placed on deriving statistical mechanisms and using the methods within data analyses. Study design and informative data summaries motivate the statistical inference techniques for linear models. Statistical thinking and communication skills are developed through analysis of data from a variety of fields. A statistical programming language is used for data visualization, analysis, and simulations.

Prerequisite

MA381 (Introduction to Probability)

Note

A student may not receive credit for both MA223 and MA382.

No textbook required

Due to the unique nature of the content within MA382, there isn't an existing text that I think is worth purchasing. If you are looking for a text reference on the statistics methods we cover, the textbook that is used in MA223 would be OK. (Devore, Farnum, Doi *Applied Statistics for Engineers and Scientists*. ISBN-13: 978-1-133-11136-8)

Hybrid Format

Face-to-Face Class Meetings

Monday and Thursday @ 8am, 9am in O233

The face-to-face class meetings will have a similar feel to a traditional lecture (e.g. instructor-led content and activities). In order to accommodate the potential need for students to quarantine/stay home due to illness, all of the lecture material from the face-to-face days will also be available electronically in the Moodle site.

Online Days

Tuesday and Friday, asynchronous

The online days will contain a mixture of new content, example data analyses, and activities. *Important: Although we will not meet at a specific time, it is expected that students work through the provided material at some point during the specified day*. The face-to-face meetings will operate under the assumption that the previous days' online material was completed.

Office Hours

All office hours will be implemented using Microsoft Teams. (This is in accordance with the Rose Ready Guide.) I have scheduled Teams meetings for each Monday and Thursday from 10-10:50am and 12:00-12:50pm. Anyone is welcome to log-on during these times. No new or additional course contents will be delivered during office hours. You are encouraged to bring specific questions about the course content or assignments.

Here is the **LINK** to the Monday, Thursday 10:00-10:50 office hour meeting.

Here is the **LINK** to the Monday, Thursday 12:00-12:50 office hour meeting.

If you would like to schedule a meeting individually or outside of these office hours, please email me *at least 24 hours in advance* so that we can schedule a meeting time that works for both of us.

Course Structure

You will find that this is NOT a typical mathematics course.

A poor statistical analysis may contain completely valid computations. However, an exemplary statistical analysis not only contains correct mathematical computation, but also, a thorough data exploration and interpretation of results in context.

Additionally, most data scientists, mathematicians, and engineers will need to implement a statistical analysis at some point in their career. Due to the breadth of the statistical field, it is unlikely you will need the exact statistical tool covered in an introductory course. We strive to give you the basic skills needed to problem-solve and seek the correct analysis technique.

Topics we will cover

Schemes for collecting data; Numerical and graphical summaries; Linear models, including simple linear regression and ANOVA; Statistical inference with confidence intervals and hypothesis testing; Parametric and nonparametric techniques for implementing statistical inference; Experimental design, including the use of blocking variables;

Computing

We will exclusively use R for data analysis in this course. Please see the Getting Started with R section.

Role of the MA381 prerequisite

We will mostly use the material related to expected value, variance, moment generating functions, and transformations of random variables. Since it may have been months or years since you took this course, I will briefly review the necessary information before expecting you to use it to complete theoretical proofs. We will also have several examples of theoretical proofs during lecture to help set expectations.

Lecture Structure

Face-to-Face

Most generally, face-to-face lectures will each a have a notes packet that is covered during the class sessions. The instructor will provide print copies of the packet (when possible) to all students attending the sessions. Students may take notes within the paper packet or on the electronic packet within Moodle. Some face-to-face lectures may contain activities or software demonstrations. It is recommended to have your laptop available for use during these sessions.

Online

The online lectures will be a mixture of content delivery, example data analyses, and simulations assignments. These are completed asynchronously. Students may need to fill in notes packets, follow along to complete an example data analysis in R, or produce a set of results using simulations in R. You may wish to print off notes packets that are provided during these sessions. Any assessments (i.e. analysis or theory and simulations) that are directly related to these sessions will be specifically referenced in the course materials.

All lecture materials will contain a set of objectives that students should be able to complete by the end of the session. These objectives are then assessed using one (or more) of our five class assignment types.

Assessment Structure

This course was created using "backwards design." This implies that lecture, activity, and assessments are created based on pre-set course goals and objectives.

Each assessment category is mapped to an overall course goal. Thus, we have five overarching goals, each of which is related to a specific assessment type. This translates to having approximately 3 required assessments each week. However, you will notice that the individual assessments are short. (For example, the Analysis and Theory & Simulations are each only one question.) There will also be several optional assessments that students may choose to complete as practice for exam short answer and in preparing for the data challenge submission.

In the following subsections, you will find a description of each assessment type and its associated goal and objectives.

Quizzes

Quiz assignments are written to align with the following course goal and objectives.

Goal I: Students will have statistical literacy.

In order to achieve this goal, students will...

> *I.a* <u>Define, recognize</u>, and <u>use</u> statistical terminology correctly and in context.

> *I.b* <u>Identify</u> and <u>describe</u> relevant qualitative statistical summaries for a dataset and question of interest.

> *I.c* <u>Describe</u> relationships among different distributions and how the distributions are related to the statistical inference method.

> *I.d* <u>List</u> assumptions associated with the employed statistical inference method.

> *I.e* <u>Describe</u> what the employed statistical inference technique is doing and how it works.

General policies related to quizzes

- There will be 8 quiz assignments, occurring weekly during weeks with no exam.
  - Each quiz will consist of 5 multiple choice, true/false, or fill in the blank questions and will be implemented using Moodle.
- Each quiz will have an allotted time of 30 minutes (although it is expected that the assignment would take approximately 10 minutes) and can be completed at any time during the week.
- The quiz must be submitted by 5pm on the final day of class for the week. (For example, when we begin the term, Tuesday is the final day of the week.)
- Generally, the quiz will contain basic material from the given week (through the 3rd day's lecture).
- Students are to work individually but can use any course materials on the assignment.
  - Students who collaborate on quiz assignments will receive no completion credit for the offense. Academic Affairs will also be notified.
- The quiz is considered to be a successful completion if *at least 4* questions are correct.
  - Students will be able to view whether a quiz was a successful completion after the 5pm submission deadline.
  - The results related to individual quiz questions will be available after all re-submissions are completed.

Students who wish to use an "Outlier" to re-submit a quiz must notify Dr. Heyman within 24 hours of the submission deadline. Dr. Heyman will email the student once the quiz attempt is re-opened and the student will then have 24 hours to complete the re-submission.

Analyses

Analysis assignments are written to align with the following goal and objectives.

Goal II: Students will know how to implement statistical analyses.

In order to achieve this goal, students will...

> **II.a. Use statistical software to create visual displays of data, calculate summary statistics, and create reproducible statistical analyses.**

> **II.b Formulate parameter(s) (and hypotheses, when appropriate) associated with a question if interest.**

> **II.c Support the use of a statistical method based upon the underlying reasonable assumptions.**

> **II.d Make a decision based on the statistical analysis result which is based on a question of interest.**

> **II.e Interpret relevant information from the statistical analysis output.**

General policies related to analyses

- There will be 8 analysis assignments, occurring weekly during weeks with no exam.
  - Each analysis will consist of 1 question which students complete using R.
  - Assignments will be collected using Moodle.
- Each assignment must be submitted by 5pm on the first lecture of the following week.
- The analysis generally will contain material from the given week.
- Each student should complete their own assignment.
  - It is permissible to work with other students. However, any of the following may be considered academic misconduct:
    - utilizing or copying another students' code
    - utilizing or copying another students' discussion
    - utilizing or copying another students' collected data
  - Students engaging in academic misconduct will receive no completion credit for the offense. Academic Affairs will also be notified.
- The analysis is considered to be a successful completion if *all criteria on the provided rubric are met.*

Students who wish to use an "Outlier" to re-submit an analysis must notify Dr. Heyman within 24 hours of receiving the graded assignment. The re-submission will be due within 72 hours of receiving the graded assignment.

Exams

Exam assignments are written to assess the following goal and objectives.

<div>

Goal III: Students will think statistically.

In order to achieve this goal, students will...

III.a. Identify the scheme used to collect a dataset and describe limitations related to the parameter(s) of interest and/or statistical inference technique.

III.b Recommend a good data collection procedure when presented with a question of interest.

III.c Interpret statistical results correctly and in context.

III.d Explain why common incorrect analysis interpretations are incorrect.

III.e Address both statistical and practical significance when discussing analysis results.

III.f Describe different types of variability that are present within a dataset or statistical modeling approach.

III.g Expound upon the randomness associated with statistical summaries (especially, how their interpretation differs from traditional deterministic mathematical models)

</div>

General policies related to exams

- There will be 2 midterms and one **final exam** within the course.
    - Each exam will consist of 10 short answer questions and will be implemented using Moodle. (Potentially, there could be bonus questions within any exam.)
- Each exam will have an allotted time of 90 minutes (although it is expected that the assignment would take approximately 50 minutes) and can be completed at any time during the 24 hour period of the implementation day.
- The exam may contain any material covered previously during the term.
- Students are to work individually but can use any course materials on the assignment.
    - Students who collaborate on an exam will receive no completion credit for the first offense *AND* cannot drop the failing exam grade from their final course score. The professor may also report the offense to Academic Affairs. The second and further offenses will also receive no credit as well as be reported to Academic Affairs.
- The exam is considered to be a successful completion if *at least 8* questions are correct.
    - Students will have access to the grading rubric which is used for each short answer question.
    - Each week, there will be an optional short answer question that students may attempt as practice for the exam questions. The weekly (optional) short answer questions must be submitted by 5pm on final day of class each week.

Data Challenge

The data challenge assignment is written to align with the following course goal and objectives.

Goal IV: Students will develop their communication skills.

In order to achieve this goal, students will...

    IV.a <u>Create</u> and <u>describe</u> aesthetic graphical displays of data.

    IV.b <u>Collaborate</u> with other students to analyze data.

    IV.c <u>Describe</u> limitations associated with a statistical analysis.

    IV.d <u>Recommend</u> new data collection approaches and improvements to the data collection process for future studies.

    IV.e <u>Suggest</u> directions for further analysis of an existing dataset.

    IV.f <u>Write</u> a formal report containing results of the statistical study.

General data challenge policies

Please see the Data Challenge section.

Theory & Simulations

The theory and simulations assignments are written to assess the following goal and objectives.

Goal V: Students will understand the theory underlying statistical methods.

In order to achieve this goal, students will...

V.a <u>Find</u> the expected value and variance of statistical estimators.

V.b <u>Derive</u> the sampling distribution or null distribution associated with a statistic of interest.

V.c <u>Use</u> simulations to confirm the derived theoretical results. Also, <u>use</u> simulations to show the extent that the theoretical results do not hold when assumptions are not true.

V.d <u>Use</u> simulations to demonstrate theoretical results which are not formally proven.

V.e <u>Substantiate</u> the connection between theoretical results and simulations.

General policies related to theory & simulations

- There will be 8 theory and simulations assignments, occurring weekly during weeks with no exam.
- Each theory & simulations will consist of 1 question which students complete either using R or by-hand (proof).
- Assignments will be collected using Moodle.
- Each assignment must be submitted by 5pm on the first lecture of the following week.
- The theory and simulations generally will contain material from the given week.
- Each student should complete their own assignment.
  - It is permissible to work with other students. However, any of the following may be considered academic misconduct:
    - utilizing or copying another students' code
    - utilizing or copying another students' proof
    - utilizing or copying another students' work
  - Students engaging in academic misconduct will receive no completion credit for the offense. Academic Affairs will also be notified.
- The theory and simulations is considered to be a successful completion if *all criteria on the provided rubric are met.*

Students who wish to use an "Outlier" to re-submit a theory and simulations assignment must notify Dr. Heyman within 24 hours of receiving the graded assignment. The re-submission will be due within 72 hours of receiving the graded assignment.

Grades

I want this course to create a strong statistical foundation for you, even if only a small foundation. **In order to help you achieve this, I will be implementing a version of "specifications-based grading."** That is, instead of taking a weighted average of points earned on a series of assignments throughout the term, you will earn grades based on the requirements you choose to complete. Each assignment is graded pass/fail (meaning no partial credit is awarded).

In order to help you achieve a passing score, very clear expectations will be provided on every assignment. You have complete control over your grade in the course; I will provide you with the necessary tools and feedback to help you achieve your desired grade.

You will also notice that I am not expecting perfection. For example, a student may achieve an A in the course without having to pass every single course assignment. However, successfully completed/passing assignments will be based on clear, high standards.

Outliers

Life happens. We all have different priorities, and at times the priorities may not easily align with the timing for class assignments. I recognize that this happens with the busy schedules that students have at RHIT and want to give every student adequate opportunity to succeed.

Every student will begin the term with 2 "Outliers." Each "Outlier" provides a second chance -- you may resubmit an assignment that did not pass the first time or that you were not able to submit by the deadline. Outliers are valid for use on quizzes, analyses, and theory & simulations assignments. The guidelines for using an Outlier are provided in each assignment description within the Course Structure section.

I'm also allowing students to re-submit any/all assignments collected during Week 1 as a way for you to adjust to the specifications grading scheme. That is, any assignment that you do not pass during week 1 may be re-submitted once for "free."

In statistics, outliers are observations/cases which are very different from the rest of the data.

Letter Criteria

Final letter grades in this course will be assigned based on the mastery of course material demonstrated throughout the term.

In order to earn a grade of "D"

Students must...
- Successfully complete/pass at least 4 analysis assignments
- Successfully complete/pass at least 4 quiz assignments
- Successfully complete/pass at least 4 theory & simulations assignments
- Achieve one of:
  - Successfully complete/pass the final data challenge
  - Receive a 60% (6/10 questions passing) on *two* course exams

In order to earn a grade of "C"

Students must...
- Successfully complete/pass at least 5 analysis assignments
- Successfully complete/pass at least 5 quiz assignments
- Successfully complete/pass at least 5 theory & simulations assignments
- Successfully complete/pass the final data challenge
- Receive a 60% (6/10 questions passing) on *two* course exams

In order to earn a grade of "B"

Students must...
- Successfully complete/pass at least 6 analysis assignments
- Successfully complete/pass at least 6 quiz assignments
- Successfully complete/pass at least 6 theory & simulations assignments
- Successfully complete/pass the final data challenge
- Receive a 70% (7/10 questions passing) on *two* course exams

In order to earn a grade of "A"

Students must...
- Successfully complete/pass at least 7 analysis assignments
- Successfully complete/pass at least 7 quiz assignments
- Successfully complete/pass at least 7 theory & simulations assignments
- Successfully complete/pass the final data challenge
- Receive a 80% (8/10 questions passing) on *two* course exams

An "F" is assigned when the criteria to earn a "D" are not achieved.

"Plus" Criteria

Students may earn a "+" on their final letter grade if either of the following cases occurs. (Exception: I can't give an A+ within the RHIT system.)

- Both "Outliers" remain/were not used during the term
- A student achieves more than the required number of passes in at least one of the quiz, analysis, or theory & simulations assignment types.

Below are a couple example scenarios to help explain the way grades are assigned.

Example 1:
Rosie passed 6 analysis assignments, 7 quizzes, and 7 theory & simulations assignments. She earned a 70% on **midterm exam 1**, 50% on **midterm exam 2**, and 60% on the **final exam**. She also passed the data challenge and has both Outliers remaining. Rosie will receive a grade of C+.

Example 2:
Rulman passed 7 analysis assignments, 8 quizzes, and 6 theory & simulations. He earned an 80% on both midterms (and then didn't have to take the final!). He also passed the data challenge but has no Outliers remaining. Rulman will receive a grade of B+

Example 3:
Namluh passed 4 assignments in each category: analyses, quizzes, and theory & simulations. Zhe earned a 40% on both midterm exams and chose not to take the **final exam**. Zhe did pass the final data challenge and had both Outliers remaining. Namluh will receive a grade of D+.

Participation

I want our classroom to be an open and safe environment to share and discuss ideas. Thus, it is expected that students and the professor are good citizens during and outside of class, treating each other with mutual respect.

- **Discussion and questions are encouraged during class**. Feel free to speak up at any time during lecture sessions. My classroom is a fairly informal atmosphere.
- Any student treating others with disrespect may incur a grade penalty at the discretion of the professor, after individual discussion.
- If anything occurs during or outside of our class which makes you feel uncomfortable, please reach out to me through email or office hours. I want all students to feel comfortable in my classroom, and communication is key.
- Lecture attendance is integral to success in this course. Many assignments occur during class and it is easy to fall behind from missing 1 or 2 class sessions.
  - With the hybrid format, it is especially important to keep on schedule, since we have fewer face-to-face meetings.
- Please do not use your cell phone during class. It is disruptive to other students and the professor. Laptop use should be responsible and appropriate to our class activities.

Rose-Hulman Institute of Technology is **committed to being an inclusive community** in which the multiplicity of values, beliefs, intellectual viewpoints, and cultural perspectives enrich learning and inform scholarship.

Institute Policies

Students with Accessibility Needs

Rose-Hulman Institute of Technology strives to make all learning experiences accessible to students. If you anticipate or experience academic barriers based on accessibility issues, please feel free to communicate those needs and register with Student Accessibility Services.

Student Accessibility Services will work with you understand the process and to determine what accommodations are most appropriate for your individual situation. Visit **the Accessibility Services website** for more information. Please note that accommodations are not retroactive and accommodations cannot be provided until verified.

 Please contact Student Accessibility Services for more information at HMU 156, 812-877-8040, or eaton1@rose-hulman.edu.

Please note that it is the student's responsibility to request any approved, documented academic accommodations (such as extra time) at least three days in advance of exams.

Academic Integrity

Plagiarism, offering, and accepting solutions to assessments when a student is required to work individually are cases of academic misconduct. These instances are taken seriously by the professor and university, and will be reported. Anyone found cheating will not be permitted to withdraw and will be (appropriately) heavily penalized in the course. Academic affairs will be informed.

The **Student Handbook** and Rose-Hulman's **Academic Rules and Procedures** describe penalties and processes invoked as a consequence if academic misconduct (such as cheating, plagiarizing, or interfering with the academic progress of other students) takes place. It is the responsibility of each student to know and follow Rose-Hulman's rules about academic integrity.

Incompletes

Incompletes are granted only when a student is forced to miss several days of class due to extraordinary circumstances such as a documented confining illness or family emergency. To receive an incomplete grade, students should obtain approval from the professor before the last day of class.

No incompletes will be given unless a prior written agreement with the professor exists, following typical Rose-Hulman policy. In this course, an "I" grade will be given only in cases of extreme hardship. Poor performance in the course is not grounds for an incomplete. An incomplete shall be recorded as an "F" if the work is not completed within the time agreed upon by the professor and student.

Emergency Information

To receive email or text messages regarding emergency situations that may impact campus and, possibly, the delivery of classes, **register for RAVE alerts** *and/or follow* **@Rose-HulmanAlert on Twitter**. *Any announcements about the Institute's ability to offer classes will be shared on* **Rose-Hulman's website**.