IST 777[1]

Statistical Methods in Information Science and Technology

Location: Hinds 120
Instructor: Kevin Crowston
Email: crowston@syr.edu

Course Catalog Description:

Classical statistical procedures used in information transfer research. Emphasis on underlying rationale for each procedure and on criteria for selecting procedures in a given research situation.

Prerequisite / Co-requisite: No prerequisite courses. However, it is expected that students understand basic algebra and graphs such as bar charts or scatterplots and are able to use a spreadsheet program and to install a program on their laptops.

Audience: The course is intended primarily for students in the PhD program, though it can be of interest and value to students in other programs.

Credits: 3 credits

Course summary and objectives:

The world contains many phenomena that are variable, complex, and interrelated. The relations among such variables are seldom straightforward: they are typically noisy, causally complex, and characterized by unpredictability. We can examine and measure aspects of the world in order to understand it. When we observe or measure phenomena, we generate so many numbers of many different kinds that it can be difficult to make sense out of them without helpful tools. Quantitative analysis provides the tools that help us understand the variability and inter-relatedness of the world as we measure it. Quantitative analysis includes statistics but is more than just statistics. Graphs, tables, and other data displays are important quantitative analysis tools. Data cleaning, screening, recoding, sorting, and filtering are also important quantitative analysis tools.

This course can help you understand how to use some of these tools. The goal is not to become a statistician, mathematician, or computer scientist, but rather to become a sophisticated user and consumer of techniques that have been developed by these specialists. Being a sophisticated user will help you with your own research efforts. Perhaps of equal importance, being a sophisticated consumer will help you understand the meaning and limitations of other people's data and the conclusions they draw from it. Numbers are misused on a daily basis both intentionally and unintentionally, particularly in fields such as this one that do not have rigid norms or orthodoxy for data collection, analysis, and presentation.

Thus the primary goal of this course is for you to learn the techniques and concepts that facilitate drawing sensible conclusions from samples of quantitative data. Upon successful completion of this course, you will have developed the following skills and knowledge:

[1] IST777 Materials – Copyright 2018 by Jeffrey Stanton – Please do not post online.

- Understanding how data collection, the data themselves, and the analysis processes relate to the kinds of inferences that can be drawn
- Knowing in advance what kinds of analysis will be feasible and developing the skill of planning data collection and measurement to facilitate appropriate analysis
- Understanding the limitations of existing data sets and how their provenance influences what analyses to perform and what inferences to draw
- Knowledge of what analysis techniques to choose, based on the conclusions you wish to draw
- Knowledge of statistical inference from a broad conceptual perspective and from the perspective of one or more contemporary philosophies
- Knowledge of assumptions that data must meet for analyses and inferences to be reasonable
- Skills in preparing data for analysis, including screening your data and dealing with missing data
- Skills in interpreting results and communicating them to others
- Supplemental knowledge of strategies for obtaining numeric measurements of various relevant phenomena

Textbook and Readings:

One textbook is required for this course: Stanton (2017), *Reasoning with Data: An Introduction to Traditional and Bayesian Statistics Using R* (abbreviated below as RwD; ISBN-13: 978-1462530267; ISBN-10: 1462530265). The paperback version of this textbook is available from the bookstore. Paperback, hard back, and electronic versions are all available on Amazon and other online book sellers.

We will use R and R Studio extensively throughout this class. You will need to have access to a laptop computer so these open source packages can be installed (they work on Mac, Windowns, and Linux). Have a laptop available during each class with R and R Studio installed and ready to run. You will also find it advantageous to have R and R Studio available when you are reading the textbook and when you are reviewing the asynchronous material.

Class structure:

The class meets weekly for three hours. Generally the class will include a brief overview of the goals of the class session, review of the most recently graded homework, an in-class exercise, followed by discussion and Q & A. Please bring a functional laptop with R and R Studio installed to each class. If your laptop has a short battery life, make sure to bring your power cable and adapter.

- **Prior to class**, read the assigned chapter in *Reasoning with Data.* Each chapter runs about 15 pages and the whole book was designed to be highly readable and accessible, even to those with a limited background in math. Some may even find it entertaining. Depending upon how quickly you read, each chapter should take you no more than about 90 minutes for a thorough consideration. You will find it advantageous to run the code examples shown in the chapters as you read.
- **Following the class, you will have 72 hours to complete the homework** assignment for that week. The homework assignment generally comprises a few problems drawn from the exercises at the conclusion of each chapter of *Reasoning with Data.* Most of the homework problems require the use of R Studio. If you tackle these problems immediately after class (or even a little before!) you should be able to complete each

homework in somewhere between one and three hours. You can ask questions by email, but you must allow 24 hours for a response, so getting started early is paramount.

Course Calendar:

Note that the schedule is subject to change, that is, it is a plan rather than a contract.

Week	Date	Goals
1	27 Aug	Topic: **Getting started**; Read RwD Introduction before class; Personal introductions; Get R and R-Studio Installed; Try R and R-Studio; Initial learning and skills assessment
2	10 Sep	Topic: **Descriptive Statistics**; Read RwD Chapter 1 before class; Explore descriptive statistics and distributions; View data sets in R; Initial skills assessment returned and discussed
3	17 Sep	Topic: **Probability**; Read RwD Chapter 2 before class; Medical test accuracy exercise; Read data into R and diagnose;
4	24 Sep	Topic: **Sampling Distributions**; Read RwD Chapter 3 before class; Generating sampling distributions exercise;
5	1 Oct	Topic: **Statistical Inference Part I**; Read RwD Chapter 4 before class;
6	8 Oct	Topic: **Statistical Inference Part II**; Read RwD Chapter 5 before class; In class "quick paper" on statistical inference
7	15 Oct	Topic: **ANOVA & Experimental Groups**; Read RwD Chapter 6 before class; In class experimental data collection and analysis
8	22 Oct	Topic: **Measures of association**; Read RwD Chapter 7 before class; In class correlation data collection and analysis
9	29 Oct	Topic: **Multiple Regression/Linear Prediction**; Read RwD Chapter 8 before class; In class analysis of data analysis from published articles
10	~~5 Nov~~[2]	Topic: **Interactions**; Read RwD Chapter 9 before class; In class "quick paper" on interpreting interactions
11	~~12 Nov~~[3]	Topic: **Categorical Analysis**; Read RwD Chapter 10 before class; In class categorical prediction exercise
12	26 Nov	Topic: **Principal Components**; In class scale development exercise
13	3 Dec	Topic: **Time Series Analysis**; Read RwD Chapter 11 before classs; In class time series exercise
14	10 Dec	Topic: **Dealing with Too Many Variables**; Read RwD Chapter 12 before class; Final "quick paper" on the selection of statistical methods for different data situation
15	17 Dec	Final examination

Student Assessment:

This course provides knowledge and practice in quantitative data analysis. There are two major components of student assessment: your weekly homework and two examinations. Additionally, your active and constructive involvement in class will benefit you and your classmates. Some discussions and activities will relate to your understanding of concepts, while others will pertain to the practice of quantitative analysis. I will assign work for you to complete

[2] This class will need to be rescheduled because of the CSCW conference.
[3] This class may need to be rescheduled because of the ASIST conference

in class, sometimes individually and sometimes in small groups. Certain classes will contain "low stakes" practice tests (that is, you get credit just for trying your best). Generally, I will assign exercises in class that stretch your skills and knowledge. In these cases, I will evaluate the extent to which you engage with the problems I present, and the extent to which you improve your own understanding of quantitative methods. Note that the primary goal of these assessments is to enhance your learning. If you work hard, jump in with both feet, and do all of the assigned work, I will consider it a huge success, and I can assure you that you will obtain a fair result at the end. Here is the breakdown of points for the course:

10 homeworks, 5 points each = 50 points
5 practice tests, 2 points each = 10 points (weeks 4, 5, 8, 9, and 10)
1 midterm w/roughly 15 questions = 15 points
1 final exam w/roughly 25 questions = 25 points

And the grading table:
A = 95-100; A- = 90-94.99; B+ = 85-89.99; B = 80-84.99; B- = 75-79.99
C+ = 70-74.99; C = 65-69.99; C- = 60-64.99; <60 = F

No rounding; no extra credit; no late submissions; no make up assignments.

According to the grading policy of the School of Information Studies, a basic grade will be awarded for student performance that is judged to be satisfactory for the course level (undergraduate or graduate). All other grades will be determined in comparison with the standards of the basic grade. For graduate students the basic grade is B. Fulfilling the requirements for an assignment coupled with the absence of errors (in writing, arithmetic, formatting) will earn a grade no higher than B+. To earn an A- or A grade, the assignment must go beyond the minimum expected in terms of quality (e.g., insight, creativity, analysis, thoroughness, synthesis).

Failure to complete any course requirement (e.g., missing a problem set) will result in a course grade of B or lower, regardless of the grades received on other components.

An incomplete grade, I, can be given only if the circumstances preventing the on-time completion of all course requirements were clearly unforeseeable and uncontrollable. If an incomplete is required a written contract must be completed that specifies the missing work, the date by which it will be completed, and the default grade that will be given if that deadline is missed.

If you wish to discuss a grade, submit a written explanation of your argument and arrange for a private conversation. Except for unusual circumstances, no appeal for an individual assignment or project will be considered more than two weeks after the graded paper is returned. For final course grades, no appeal will be considered more than two months after the final day of classes.

It is unfair to allow some students additional opportunities, such as extra credit assignments, without allowing the same options to all students. Accordingly, extra credit assignments are not possible.

Preparing Your Homework for Submission

Prepare your responses to the assigned questions with a word processor, cutting and pasting output from R Studio as appropriate. In general, if you are providing R output, you should include the snippet of code that created that output. When R provides output that is formatted as a table, you may find it helpful to switch to a monospace font such as Courier.

Submit your homework as a PDF file. Name your file HWX_Lastname.pdf, substituting the week number and your own last name. The homework intentionally models the kinds of actions and language you may use as a researcher, so it is important to format the homework in a professional manner. Submitting a PDF ensures that the way you submit the homework is the way it will appear when graded.

The main purpose of the homework is to practice the skills you have learned that week and cement the knowledge that you have gained. As such, **homework should be a solo activity**, so that you can prove to yourself and the instructor your capacity to accomplish the work independently. To the extent that you do collaborate with someone else – including seeking coaching, feedback, suggestions, or code examples – **you must acknowledge this at the top of the homework file**. This is the "give credit where credit is due" principle and it is paramount for scientists. The same idea holds with respect to consulting outside resources, such as the R-Bloggers website. Based on these principles, your homework should begin with a statement like one of these:

> Homework 1 by Fred Flintstone: I produced the material below with no assistance.
>
> or
>
> Homework 1 by Fred Flintstone: I consulted with Barney Rubble about how to tackle these problems, but we each wrote our code and text independently.

Any variation on the second statement is reasonable as long as it is forthright. Once you submit a homework that reports collaboration, your instructor may have new advice, guidance, or suggestions for you to enhance your learning.

Course Specific Policies:

Conduct of discussions. It is expected that students will behave professionally both in language and attitude when commenting or responding to discussions. Public disparagement of your fellow students in this course is unacceptable and may result in disciplinary action. Additionally, discussions should model your ability to think critically about course topics and articulate ideas clearly. Responses should be detailed and explanatory. Simple unconstructive replies such as "I agree (or disagree) with the previous comment" will not be considered valid unless elaborated upon. If you have any concerns regarding the suitability of a comment (yours or another's), please contact the instructor.

Guidelines for citations. When you prepare assignments or post on the discussion boards be sure to provide proper bibliographical information for any sources referenced, for direct quotations and for the source of key concepts or ideas. It is critical to include quotation marks and citations when you cut and paste from any source. Any citation format is acceptable (I personally use APA format), as long as it provides sufficient information for a reader to find the source (i.e., authors names, title of article or book, title, volume and issue of journal (if appropriate), page numbers, publisher, date of publication). If you cite a webpage, be sure to indicate the URL and the date on which you accessed the page, as pages do change. Failure to cite sources is considered plagiarism and subject to sanctions ranging from being required to redo the assignment through expulsion (see above). If you have any questions about what must be cited or how to cite, please feel free to ask.

Writing. In addition to punctuality, grammar, presentation and ability to follow instructions are very important, as in the real world. If your work does not meet professional standards, up to

30% of your score may be deducted. It is essential that you spell check and proofread your documents.

In assignments you may be asked to offer opinions. As with discussions, simply stating your opinion does not constitute a complete response. You must support any opinion with arguments and evidence.

University Policies:

Students should review the University's policies regarding: *Diversity and Disability* https://www.syracuse.edu/life/accessibilitydiversity/; the *Religious Observances Notification and Policy* http://supolicies.syr.edu/studs/religious_observance.htm; and *Orange SUccess* http://orangesuccess.syr.edu/getting-started-2/.

University Attendance Policy

Attendance in classes is expected in all courses at Syracuse University. Students are expected to arrive on campus in time to attend the first meeting of all classes for which they are registered. Students who do not attend classes starting with the first scheduled meeting may be academically withdrawn as not making progress toward degree by failure to attend. Instructors set course-specific policies for absences from scheduled class meetings in their syllabi.

It is a federal requirement that students who do not attend or cease to attend a class to be reported at the time of determination by the faculty. Students should also review the university's religious observance policy and make the required arrangements at the beginning of each semester.

Academic Integrity Policy

Syracuse University's Academic Integrity Policy reflects the high value that we, as a university community, place on honesty in academic work. The policy defines our expectations for academic honesty and holds students accountable for the integrity of all work they submit. Students should understand that it is their responsibility to learn about course-specific expectations, as well as about university-wide academic integrity expectations. The policy governs appropriate citation and use of sources, the integrity of work submitted in exams and assignments, and the veracity of signatures on attendance sheets and other verification of participation in class activities. The policy also prohibits students from submitting the same work in more than one class without receiving written authorization in advance from both instructors. Under the policy, students found in violation are subject to grade sanctions determined by the course instructor and non-grade sanctions determined by the School or College where the course is offered as described in the Violation and Sanction Classification Rubric. SU students are required to read an online summary of the University's academic integrity expectations and provide an electronic signature agreeing to abide by them twice a year during pre-term check-in on MySlice. For more information about the policy, see http://class.syr.edu/academic-integrity/policy/.

For IST 777, you are permitted and encouraged to discuss course material with your classmates. However, individual assignments should not be worked on in a collaborative manner, as noted above.

The Violation and Sanction Classification Rubric establishes recommended guidelines for the determination of grade penalties by faculty and instructors, while also giving them discretion to select the grade penalty they believe most suitable, including course failure, regardless of

violation level. Any established violation in this course may result in course failure regardless of violation level.

Disability-Related Accommodations

Syracuse University values diversity and inclusion; we are committed to a climate of mutual respect and full participation. If you believe that you need academic adjustments (accommodations) for a disability, please contact the Office of Disability Services (ODS), visit the ODS website– http://disabilityservices.syr.edu, located in Room 309 of 804 University Avenue, or call (315) 443-4498 or TDD: (315) 443-1371 for an appointment to discuss your needs and the process for requesting academic adjustments. ODS is responsible for coordinating disability-related academic adjustments and will issue students with documented Disabilities Accommodation Authorization Letters, as appropriate. Since academic adjustments may require early planning and generally are not provided retroactively, please contact ODS as soon as possible. Our goal at the iSchool is to create learning environments that are useable, equitable, inclusive and welcoming. If there are aspects of the instruction or design of this course that result in barriers to your inclusion or accurate assessment or achievement, please meet with me to discuss additional strategies beyond official accommodations that may be helpful to your success.

Discrimination or Harassment

The University does not discriminate and prohibits harassment or discrimination related to any protected category including creed, ethnicity, citizenship, sexual orientation, national origin, sex, gender, pregnancy, disability, marital status, age, race, color, veteran status, military status, religion, sexual orientation, domestic violence status, genetic information, gender identity, gender expression or perceived gender.

Any complaint of discrimination or harassment related to any of these protected bases should be reported to Sheila Johnson-Willis, the University's Chief Equal Opportunity & Title IX Officer. She is responsible for coordinating compliance efforts under various laws including Titles VI, VII, IX and Section 504 of the Rehabilitation Act. She can be contacted at Equal Opportunity, Inclusion, and Resolution Services, 005 Steele Hall, Syracuse University, Syracuse, NY 13244-1120; by email: titleix@syr.edu; or by telephone: 315-443-0211.

Federal and state law, and University policy prohibit discrimination and harassment based on sex or gender (including sexual harassment, sexual assault, domestic/dating violence, stalking, sexual exploitation, and retaliation). If a student has been harassed or assaulted, they can obtain confidential counseling support, 24-hours a day, 7 days a week, from the Sexual and Relationship Violence Response Team at the Counseling Center (315-443-4715, 200 Walnut Place, Syracuse, New York 13244-5040). Incidents of sexual violence or harassment can be reported non-confidentially to the University's Title IX Officer (Sheila Johnson Willis, 315-443-0211, titleix@syr.edu, 005 Steele Hall). Reports to law enforcement can be made to the University's Department of Public Safety (315-443-2224, 005 Sims Hall), the Syracuse Police Department (511 South State Street, Syracuse, New York, 911 in case of emergency or 315-435-3016 to speak with the Abused Persons Unit), or the State Police (844-845-7269). I will seek to keep information you share with me private to the greatest extent possible, but as a professor I have mandatory reporting responsibilities to share information regarding sexual misconduct, harassment, and crimes I learn about to help make our campus a safer place for all.

Religious Observances Notification and Policy

SU's religious observances policy, found at http://supolicies.syr.edu/emp_ben/religious_observance.htm , recognizes the diversity of faiths represented in the campus community and protects the rights of students, faculty, and staff to observe religious holy days according to their tradition. Under the policy, students should have an opportunity to make up any examination, study, or work requirements that may be missed due to a religious observance provided they notify their instructors no later than the end of the second week of classes for regular session classes and by the submission deadline for flexible-formatted classes. Student deadlines are posted in MySlice under **Student Services/Enrollment/My Religious Observances/Add a Notification**.

Student Academic Work Policy

Student work prepared for University courses in any media may be used for educational purposes, if the course syllabus makes clear that such use may occur. You grant permission to have your work used in this manner by registering for, and by continuing to be enrolled in, courses where such use of student work is announced in the course syllabus.

I intend to use academic work that you complete this semester for educational purposes in this course during this semester. Your registration and continued enrollment constitute your permission.

Course Evaluations

There will be an end of course evaluation for you to complete this term. This evaluation will be conducted online and is entirely anonymous. You will receive a notification from the Syracuse University Office of Institutional Research & Assessment (OIRA) department in your email account with the evaluation website link and your passcode. Please take the time and fill out this evaluation as your feedback and support of this assessment effort is very much appreciated. The school carefully reviews ratings and comments that you submit, and these factor into decisions about course, program and instructor development.

End of semester evaluation will be available for completion prior to your final exams week. This evaluation is slightly longer and it is used to gauge the instructor performance and make adjustments to the course to ensure it meets our student needs.

We faculty work hard to do the best possible job when preparing and delivering courses for our students. Please understand that not only does the school use the course evaluations to make decisions about the curriculum to improve where necessary, but they also use them to make decisions about faculty members.