UT-Austin iSchool Syllabus I306 Statistics for Informatics Spring 2024

2024-01-18

Description (from the catalog)

Restricted to informatics majors and students pursuing the informatics minor. Examine fundamental principles of probability and statistics. Cultivate an understanding of descriptive and inferential statistics. Conduct and interpret statistical analyses using statistical analysis software, and apply these analyses to common issues in informatics. Three lecture hours a week for one semester. Offered on the letter-grade basis only.

Details

Important note: The information presented in this syllabus is subject to expansion, contraction, change, or stasis during the semester. In case of conflict between versions, the copy on Canvas takes precedence.

Course Number

27395

Prerequisites

none, but knowing math and programming will help

Time

MW 1700–1830

Place

RLP 0.108

Dates

January 16–April 29, 2024

Final Exam

Take-home, due at our official exam time

Instructor

Mick McQuaid

Email

mcq@utexas.edu

Office

1616 Guadalupe St, Room 5.402

Office Hours

M,W 4PM to 5PM in RLP 0.108, or by appointment at my UTA office or by appointment on Zoom at https://utexas.zoom.us/my/mickmcquaid

Materials

Our textbook is the freely available Diez, Çetinkaya-Rundel, and Barr (2019). Additional textbooks you can use include the freely downloadable Wickham, Çetinkaya-Rundel, and Grolemund (2023) and James et al. (2021). A more advanced textbook is the freely available Kuhn and Silge (2022), the full text of which is available at tmwr. You will also require the study guide, available at https://mickmcquaid.com/stats/studyGuide

Learning Outcomes

- learn to describe data using statistics and contingency tables to summarize
- learn to use probability distributions
- learn to visualize data
- learn to develop confidence intervals
- learn to conduct hypothesis tests
- learn to conduct single and multiple regression and logistic regression
- learn to write reproducible reports

Class Format

The class will be half lecture, half work time on a computer. You should definitely bring a laptop computer to class every day.

Schedule

Week 1 (17 Jan) Introduction to data — Introduction to R — Introduction to R Studio — Introduction to R markdown

Week 2 (22 Jan, 24 Jan) Summarizing data — Examining numerical data — Considering categorical data — More on R markdown — Introduction to Quarto

Week 3 (29 Jan, 31 Jan) More on R and quarto — Visually summarizing data

Week 4 (05 Feb, 07 Feb) Probability — Defining probability — Conditional probability — Sampling from a small population — Random variables — Continuous distributions — Milestone 1 due

Week 5 (12 Feb, 14 Feb) Distributions of random variables — Normal distribution — Geometric distribution — Binomial distribution — Negative binomial distribution — Poisson distribution — Milestone 2 due

Week 6 (19 Feb, 21 Feb) Foundations for inference — Point estimates and sampling variability — Confidence intervals for a proportion — Hypothesis testing for a proportion

Week 7 (26 Feb, 28 Feb) Inference for categorical data — Inference for a single proportion — Differences of two proportions — Testing for goodness of fit using chi-square — Testing for independence in two-way tables

Week 8 (04 Mar, 06 Mar) Inference for numerical data — One-sample means with the t-distribution — Paired data — Difference of two means — Power calculations for a difference of means — Comparing many means with ANOVA

Spring Break!

Week 9 (18 Mar, 20 Mar) *Note: class is cancelled on 18 Mar* — Introduction to linear regression — Fitting a line, residuals, and correlation — Least squares regression — Types of outliers in linear regression — Inference for linear regression

Week 10 (25 Mar, 27 Mar) Multiple and logistic regression — Introduction to multiple regression — Model selection — Checking model conditions using graphics — Multiple regression case study: Mario Kart — Introduction to logistic regression — Milestone 3 due

Week 11 (01 Apr, 03 Apr) Regression diagnostics

Week 12 (08 Apr, 10 Apr) More on R and the tidyverse

Week 13 (15 Apr, 17 Apr) Typical UX experiments

Week 14 (22 Apr, 24 Apr) Typical data science experiments — Milestone 4 due

Week 15 (29 Apr) Summary

Grading

I plan to grade assignments within two weeks of their due date except where circumstances interfere. The grading scale used along with the grade components follow.

Table 1: Scores are not rounded

letter grade		lower bound			upper bound
A	>=	94.0%			
A-	>=	90.0%	&	<	94%
B+	>=	87.0%	&	<	90%
B	>=	83.0%	&	<	87%
B-	>=	80.0%	&	<	83%
C+	>=	77.0%	&	<	80%

letter grade		lower bound			upper bound
C	>=	73.0%	&	<	77%
C-	>=	70.0%	&	<	73%
D	>=	60.0%	&	<	70%
F				<	60%

Note that all work requires the submission of a .qmd file and a .html file. Omission of either will result in no credit for the work.

Project work: 4 milestones, first three each 12 points, fourth is 14 points

- Milestone 1: description (tables, summary stats)
- Milestone 2: description (visualization)
- Milestone 3: regression
- Milestone 4: regression diagnostics and final report

Weekly work

- Weekly exercises, weeks 2 through 14 (2 points each for the first nine, 4 points each for the next four, 6 points for the last one)
- Attendance (10 points)

Attendance

I will take attendance every day and your percentage of attendance will count for ten percent of your grade.

If you have a legitimate need for absence, such as illness or job interview, notify the instructor by email as soon as possible and you may receive an excused absence.

POLICIES

Important Note: The policies of the University are undergoing change. The following *may* be superseded by the policies at https://utexas.instructure.com/courses/1377522, which is a Canvas course containing the honor code which you *must* adhere to, as well as much of the following information. A better URL may be https://utexas.instructure.com/enroll/TP964H if for some reason you are not enrolled in the site.

Attendance

All concerns about attendance recording must be resolved within 72 hours of the class session in question. In other words, you can't come to the instructor weeks later and insist you were present on such-and-such a day.

Assignment Submission

All assignments must be submitted via Canvas. No assignment should be submitted via email. Any assignment submitted via email will receive a grade of zero. It may be tempting to try to submit assignments via email when you have trouble with Canvas but the correct response is to contact tech support and resolve the problem with Canvas.

Extra credit and grade rounding

There is no extra credit available in this class and grades are not rounded. You receive exactly the letter grade corresponding to the score you achieve.

Disability and Access

The university is committed to creating an accessible and inclusive learning environment consistent with university policy and federal and state law. Please let me know if you experience any barriers to learning so I can work with you to ensure you have equal opportunity to participate fully in this course. If you are a student with a disability, or think you may have a disability, and need accommodations please contact Disability and Access (D&A). Please refer to D&A's website for contact and more information: http://community.utexas.edu/disability/. If you are already registered with D&A , please deliver your Accommodation Letter to me as early as possible in the semester so we can discuss your approved accommodations and needs in this course.

Policy on Academic Integrity

Students who violate University rules on academic misconduct are subject to the student conduct process and potential disciplinary action. A student found responsible for academic misconduct may be assigned both a status sanction and a grade impact for the course. The grade impact could range from a zero on the assignment in question up to a failing grade in the course. A status sanction can range from probation, deferred suspension and/or dismissal from the University. To learn more about academic integrity standards, tips for avoiding a potential academic misconduct violation, and the overall

conduct process, please visit the Student Conduct and Academic Integrity website at: http://deanofstudents.utexas.edu/conduct.

Class Recordings

Class recordings are reserved only for students in this class for educational purposes and are protected under FERPA. The recordings should not be shared outside the class in any form. Violation of this restriction by a student could lead to Student Misconduct proceedings.

Artificial intelligence

The creation of artificial intelligence tools for widespread use is an exciting innovation. These tools have both appropriate and inappropriate uses in classwork. The use of artificial intelligence tools (such as ChatGPT) in this class is permitted but must be documented. Usually, you should include a lengthy disclaimer at the end of the assignment as a separate paragraph telling which generative AI tool was used, e.g., ChatGPT, and what it was used for. Failure to document will be considered a cheating offense, punishable under the rules for academic integrity. The disclaimer must be specific and thorough. A brief, vague statement will not be considered sufficient.

Personal Pronouns

Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, culture, religion, politics, sexual orientation, gender identity & expression, and nationalities. Class rosters are provided to the instructor with the student's legal name, unless they have added a "chosen name" with the registrar's office, which you can do so here: https://utdirect.utexas.edu/apps/ais/chosen_name/. I will gladly honor your request to address you by a name that is different from what appears on the official roster, and by the pronouns you use (she/he/they/ze, etc). Please advise me of any changes early in the semester so that I may make appropriate updates to my records. For instructions on how to add your pronouns to Canvas, visit https://utexas.instructure.com/courses/633028/pages/profile-pronouns. More resources are available on the Women's Community Center website, https://community.utexas.edu/wcc/.

Basic Needs Security

Any student who faces challenges securing their food or housing and believes this may affect their performance in the course is urged to contact the Dean of Students for support.

UT maintains the UT Outpost (https://deanofstudents.utexas.edu/emergency/utoutpost.php) which is a free on-campus food pantry and career closet. Furthermore, please notify the professor if you are comfortable in doing so. This will enable him to provide any resources that he may possess.

Mental Health Information

Students who are struggling for any reason and who believe that it might impact their performance in the course are urged to reach out to Bryce Moffett if they feel comfortable. This will allow her to provide any resources or accommodations that she can. If immediate mental health assistance is needed, call the Counseling and Mental Health Center (CMHC) at 512-471-3515 or you may also contact Bryce Moffett, LCSW (iSchool CARE counselor) at 512-232-4449. Bryce's office is located in FAC18S and she holds drop in Office Hours on Wednesday from 2-3pm. For urgent mental health concerns, please contact the CMHC 24/7 Crisis Line at 512-471-2255.

Carrying of Handguns on Campus

Students in this class should be aware of the following university policies related to Texas' Open Carry Law: Students in this class who hold a license to carry are asked to review the university policy regarding campus carry.

- Individuals who hold a license to carry are eligible to carry a concealed handgun on campus, including in most outdoor areas, buildings and spaces that are accessible to the public, and in classrooms.
- It is the responsibility of concealed-carry license holders to carry their handguns on or about their person at all times while on campus. Open carry is NOT permitted, meaning that a license holder may not carry a partially or wholly visible handgun on campus premises or on any university driveway, street, sidewalk or walkway, parking lot, parking garage, or other parking area.
- Per my right, I prohibit carrying of handguns in my personal office. Note that this information will also be conveyed to all students verbally during the first week of class. This written notice is intended to reinforce the verbal notification, and is not a "legally effective" means of notification in its own right.

LGBTQIA+ Community

As an institution committed to creating a safe and inclusive learning environment, The University of Texas at Austin strictly prohibits discrimination, harassment, or marginalization based on sexual orientation or gender identity under Title IX. If you encounter any discrimination or harassment, please seek support from the Title IX office.

Additionally, we encourage you to complete the Campus Climate Survey by following the link: https://app.smartsheet.com/b/form/d70ce9db84a3403ab00394e4617f8f3b

If you experience any form of discrimination or harassment, please contact the Title IX office for support. If you do not wish to contact the UT Title IX office, you may view confidential community resources at https://titleix.utexas.edu/community-resources-confidential . The Womens Community Center, found at https://community.utexas.edu/wcc/ , offers resources and support for LGBTQIA+ students, and I encourage you to visit their website for more information or to contact their professional staff.

I am committed to creating a safe and inclusive learning environment for all students. This includes fostering an environment of respect, openness, and understanding in the classroom and actively working to address any discrimination or harassment that may occur. If you wish to display your pronouns on your Canvas page, you can find a guide here: https://utexas.instructure.com/courses/633028/pages/profile-pronouns. Furthermore, you can include a "preferred name" by viewing the following link to class rosters, which come with the student's legal name (unless an addition of a preferred name is made): https://utdirect.utexas.edu/apps/ais/chosen_name/.

TITLE IX DISCLOSURE

Beginning January 1, 2020, Texas Education Code, Section 51.252 (formerly known as Senate Bill 212) requires all employees of Texas universities, including faculty, to report to the Title IX Office any information regarding incidents of sexual harassment, sexual assault, dating violence, or stalking that is disclosed to them. Texas law requires that all employees who witness or receive information about incidents of this type (including, but not limited to, written forms, applications, one-on-one conversations, class assignments, class discussions, or third-party reports) must report it to the Title IX Coordinator. Before talking with me, or with any faculty or staff member about a Title IX-related incident, please remember that I will be required to report this information.

Although graduate teaching and research assistants are not subject to Texas Education Code, Section 51.252, they are mandatory reporters under federal Title IX regulations and are required to report a wide range of behaviors we refer to as sexual misconduct, including the types of misconduct covered under Texas Education Code, Section 51.252. Title IX of the Education Amendments of 1972 is a federal civil rights law that prohibits discrimination on the basis of sex – including pregnancy and parental status – in educational programs and activities. The Title IX Office has developed supportive ways and compiled campus resources to support all impacted by a Title IX matter.

If you would like to speak with a case manager, who can provide support, resources, or academic accommodations, in the Title IX Office, please email: supportandresources@austin.utexas.edu. Case managers can also provide support, resources, and accommodations for pregnant, nursing, and parenting students.

For more information about reporting options and resources, please visit: https://titleix.utexas.edu, contact the Title IX Office via email at: titleix@austin.utexas.edu, or call 512-471-0419.

References

Diez, David, Mine Çetinkaya-Rundel, and Cristopher D Barr. 2019. *OpenIntro Statistics, Fourth Edition*. self-published. https://openintro.org/os.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2021. *An Introduction to Statistical Learning*. Springer New York.

Kuhn, Max, and Julia Silge. 2022. *Tidy Modeling with R*. Sebastopol, CA: O'Reilly.

Wickham, Hadley, Mine Çetinkaya-Rundel, and Garrett Grolemund. 2023. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. 2nd ed. O'Reilly Media, Inc. https://r4ds.hadley.nz/.