CSci 502: Statistics for Computational Science and Analysis

Course Syllabus

Summer 2017

Instructor

> Derek Harter, Ph.D., Associate Professor
> Department of Computer Science
> Texas A&M University - Commerce
> Office: SCI 355
> Office Hours: T, Th 1 - 3 pm (or by appointment)
> E-mail: Derek.Harter@tamuc.edu

Class Meetings

This face-to-face summer class is what is known as a W10 extended summer session course. It will meet for 10 weeks over both the summer I and II sessions, from June 5, 2017 through August 10, 2017. We will be covering all of the same materials and assignments given during a normal 16 week semester session. You need to plan to schedule about 18 to 20 hours of out of class time study and assignment preparation time for this Summer W10 session. This is especially important in this class as we will be doing a lot of exercises from the textbook for assignments and class evaluations. Doing and understanding the assignments are essential for grasping the materials in this class.

0XE   40311   Meets 6/5/2017 through 8/10/2010   TR 10-11:50a Jour 110

Course Description

This course provides an introductory framework for the statistical background required for scientific computation and data analysis. The course introduces fundamental statistical concepts such as probability, random variables, probability distributions, statistical expectation, sampling distributions, hypothesis testing, linear regression, correlation, and visualization/plotting of data, with emphasis on applications to scientific computing and computational science problems. Concepts will be reinforced by having students use a statistical/scientific computing & visualization software in order to apply the concepts that they learn by solving problems from various disciplines. **Credit hours: 3**. **You have to obtain at least a B grade in order to pass this course since this is a prerequisite course for the MS CPSI program.**

Prerequisites

No prerequisites, but you need to have passed an undergraduate based mathematical sequence for the sciences in order to do well in this course. This course is a calculus and discrete based mathematical approach to developing an understanding of probability and statistics. A solid background and comfort with basic calculus and concepts from discrete mathematics greatly enhance understanding and prospects of success in this course.

Student Learning Outcomes:

Students will be able to

- (SLO #1) demonstrate understanding of the probability, random variables and probability distributions

- (SLO #2) demonstrate understanding of hypothesis testing and inference

- (SLO #3) demonstrate understanding of linear regression and correlation

- (SLO #4) demonstrate understanding of using statistical descriptors of data for analysis and visualization

- (SLO #5) demonstrate the ability to use a statistical analysis toolbox/software and apply to real data for statistical analysis and visualization.

Textbook

Required:

Probability and Statistics for Engineers and Scientists, 9th Ed. by Walpole, Myers, Myers and Ye, Prentice Hall. ISBN-13: 978-0-321-62911-1

Recommended

Probability and Statistics with R for Engineers and Scientists, 1st Ed. by Akritas, Prentice Hall. ISBN-13: 978-0321852991 http://sites.stat.psu.edu/~mga/401/

R for Data Science: Import, Tidy, Transform, Visualize, and Model Data, 1st Ed. by Wickham, Grolemund, O'Reilley Media. ISBN-13: 978-1491910399 http://r4ds.had.co.nz/

Software Carpentry Programming with R & R for Reproducible Scientific Research Lessons https://software-carpentry.org/lessons/

Course Outline / Content

W	Dte	Topic / Activity	Assg	Test
1	6/5	Ch 1 Introduction to Statisticss		
2	6/12	Ch 2 Probability		
3	6/19	Ch 3 Random Variables and Prob Distributions		
4	6/26	Ch 4 Mathematical Expectation		
5	7/3	Ch 5 Some Discrete Probability Distributions		Midterm
6	7/10	Ch 6 Some Continuous Probability Dists		
7	7/17	Ch 8 Fundamental Sampling Distributions		
8	7/24	Ch 9 One- and Two-Sample Estimation Problems		
9	7/31	Ch 10 One- and Two-Sample Tests of Hypothesis		
10	8/7	Ch 11 Simple Linear Regression and Correlation		Final

All dates and content may be subject to change throughout the semester. Changes will be communicated to the students in class or via eCollege course shell announcements or via email.

Evaluation (Tentative)

Your grade for the course will be based on the following (approximate) percentages:

Homework Assignments & Projects	60%
Midterm Exam	15%
Final Exam (comprehensive of all materials)	25%

Final Average	Letter Grade
90 - 100	A
80 - 89	B
70 - 79	C
60 - 69	D
Below 60	F

Course Requirements

For a normal 16 week course, we usually recommend you plan for at least 12 hours of study time per week for a 3 credit hour course (which will vary depending on your background and ability). We cover the same material in summer as in a regular semester, but in less than 2/3 of the total time. Thus a rough guidelines of 18 to 20 hours a week allocated for reading, studying and performing assignments should be expected for students with a required background for the course materials.

Assignments: There will be regularly assigned homework problems. These assignments may require the application of various software packages. Assignments will be given and returned via the online eCollege system as a convenience to the students and the instructor. It is the student's responsibility to login and check the course eCollege site daily for announcements, assignments and course-related content. It is very important that students follow the instructions carefully on the assignments. It is the student's responsibility to have all assignments ready on time by the given due date. Late assignment may not be accepted or may be penalized and assignment may not be accepted beyond a certain time. Important material from the text and outside sources will be covered in class. Students should plan to take careful notes as not all material can be found in the texts or readings. End of chapter activities and online activities may be assigned to reinforce material in the text.

Exams: Two exams will be given, one midterm exam and one final exam. The exams will be closed book/notes and will test assigned readings and

material discussed in class. The instructor may add other necessary exams if he sees necessary. Cellphones and other telecommunication electronics will not be allowed during the exams.

Attendance: Regular attendance is recommended in order to best grasp the materials in this course. You will be responsible for all assignments and tests given in class that you miss if you are not in attendance.

Projects: In some of the assignments, there will be applied statistical analysis projects in which the concepts will be reinforced by having students learn to use a statistical /scientific computing & visualization software and apply it to analyze data from various disciplines.

Students can see their graded assignment, quiz, exam papers, project reports and ask their questions during the office hours. The students have maximum one week to see their graded papers after the grades are announced (announced in class or uploaded to eCollege); beyond that, at the instructor's discretion. The overall course grades are finalized after all the exams, assignments, quizzes and attendances are weighed & evaluated at the end of the semester on the instructor's excel spreadsheet.

Course Deadlines

Credit will be given for ONLY those exam(s), program(s), and/or project(s) turned in no later than the deadline(s) as announced by the instructor of this class unless prior arrangement has been made with the instructor. For a 10 week course we have tight deadlines in terms of assignments and course work, so responding to and meeting deadlines is critical for assignments in this summer session, as it may not be possible to evaluate and return any work submitted after the official deadlines.

Academic Ethics

"All students enrolled at the University shall follow the tenets of common decency and acceptable behavior conducive to a positive learning environment." (See Student's Guide Handbook, Policies and Procedures, Conduct). Ethics also includes the issue of plagiarism, and copying code for programming/lab assignments is just as serious as any other type of plagiarism. If you are caught sharing or using other people's work in this class, you will receive a 0 grade and a warning on the first instance. A subsequent instance will result in receiving an F grade for the course, and possible disciplinary proceedings.

Student's with Disabilities

The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. If you have a disability requiring an accommodation, please contact: Office of Student Disability Resources and Services, Texas A&M University-Commerce, Gee Library, Room 132, Phone (903) 886-5150, StudentDisabilityServices@tamuc.edu