Statistics 3201: Introduction to Probability for Data Analytics
Autumn 2017 Course Syllabus

Instructor: Dr. O. Chkrebtii

Lectures: 1:50pm – 2:45pm on Mondays, Wednesdays, Fridays in Smith Lab 2150.

Office Hours: Monday, 9:00 am - 10:00 am, 323 Cockins Hall (CH). Individual appointments should be requested via email and will not be available on short notice. The Teaching Assistant will also be available to answer questions (see below).

Email: chkrebtii.1@osu.edu. Please **begin subject with "STAT 3201"**. To protect your privacy, **all course email correspondence must be conducted using your valid OSU name.# email account**. Due to the large volume of emails, I may not be able to answer promptly, therefore please consider whether your question would be best answered in person during office hours.

Teaching Assistant: Amartya Ghosh (ghosh.147@osu.edu). The Teaching Assistant will be available to answer questions in the Tutor Room (CH 122) during the following times of the week: Tuesday 4:10 pm – 5:10 pm and 5:20 pm – 6:20 pm; Wednesday 5:20 pm - - 6:20 pm; and Friday 9:10 am – 10:10am

Course Description: An introduction to probability and its role in statistical methods for data analytics. Equal emphasis is placed on analytical and simulation-based methods for quantifying uncertainty. Approaches to assessing the accuracy of simulation methods are discussed. Students should have some prior knowledge of basic programming. Applications of probability and sampling to big-data settings are discussed.

Upon successful completion of the course, students will be able to

1. Quantify uncertainty about events using mathematical descriptions of probability
2. Quantify uncertainty about events using simulation methods
3. Assess the quality and accuracy of simulation-based descriptions of uncertainty
4. Update a description of uncertainty based on new information
5. Identify appropriate probability models for experiments/data and summarize expected outcomes from such models
6. Use correlation and conditional expectation to describe the relationship between two random variables.

7. Quantify uncertainty about summary statistics for large data sets

Prerequisites: Prerequisites: Math 1152 or 1161.xx or 1172 or 1181 or equivalent, and CSE Placement Level A or equivalent; or permission of the instructor.

Course Website: Important announcements, course materials, homework problems, computing references, and other information about the class are posted on Carmen (carmen.osu.edu, login with your web ID).

Textbook: The required textbook for this course is:

- *Mathematical Statistics with Applications* (7th edition) by Wackerly, Mendenhall and Sheaffer

Course material will be supplemented with the freely available textbook

- *Introduction to Probability and Statistics using R* by Kerns; available online at https://cran.r-project.org/web/packages/IPSUR/vignettes/IPSUR.pdf

Course Materials: A reading list from the required textbook will be provided for each lecture. Please read the sections of the textbook that will be covered, and any additional material posted on Carmen prior to the class.

Homework Assignments: Homework will be assigned weekly or biweekly. It will consist of mostly textbook problems, problems motivated by data analytics applications, and small computer simulation problems. **Questions numbers referenced in the homework are from the textbook edition listed above: if you are using a different edition/version of the textbook, it is your responsibility to check that you have solved the correct questions. No points will be awarded from answering a question other than the one being assigned.** You are encouraged to work together on the problems, but each student must hand in his or her own work, written in his or her own words. **Do not copy any part of another student's homework including computer output**. Use of homework solutions distributed in previous offerings of the course or available on the web constitutes academic misconduct and will be handled according to university rules. **Sharing or disseminating solutions, or in any way knowingly enabling others to commit academic misconduct also constitutes academic misconduct, and will be reported.** A hard copy of the homework solutions should be submitted at the beginning of class on the due date. The solutions may be handwritten or typed. Please be sure that the questions are clearly labeled, all supporting work (including computer code) can be easily identified, and that all figures/tables are referenced and interpreted in the text. Electronic versions of homework solutions will not be accepted unless permission from the instructor is obtained in advance.

Exams: There will be three in-class exams. Coverage includes lecture material, assigned reading, and homework. Tentative dates are provided on the weekly lesson plan. Statistical tables will be provided as needed. Calculators may be used, but no communication devices are allowed (e.g. mobile phones). You may use one 8.5×11 inch handwritten sheet of paper (both sides) with formulas. Makeup exams require a valid excuse and official proof if I am notified in advance or as soon as possible. A make-up exam must be taken within a week of the missed exam. Exceptions to this policy are permitted only in extreme situations such as serious injury immediately prior to an exam or severe illness requiring hospitalization. **Note that sharing or disseminating solutions, or**

in any way knowingly enabling others to commit academic misconduct also constitutes academic misconduct, and will be reported.

Project: A class project will consist of a combination of a written portion and oral presentation, and will require use of the R software. Suggested project topics will be provided in class.

Computing: We will be using the R statistical computing software. R may be downloaded for free from http://www.r-project.org/. Many students prefer to use the interface RStudio, available for free at http://www.rstudio.com.

Attendance: Regular attendance and class participation is required. Please let the instructor know via email if you plan to miss several lectures.

Grading: In order to obtain full credit on homework and exam problems you need to show your justification for or full work. Answers without work will not receive full credit. The following is a breakdown of the final course grade:

Homework: 20%
Exam 1: 20%
Exam 2: 20%
Final Exam: 30%
Project: 10%

The following rubric will be used to compute the final letter grade: A: 93 – 100, A-: 90 – 92.9, B+: 87– 89.9: B: 83 – 86.9, B-: 80-82.9, C+: 77-79.9, C: 73 – 76.9, C-: 70–72.9, D+: 67 – 69.9, D: 60–66.9, E: below 60. The instructor reserves the right to make appropriate changes to the above if necessary. However, as usual there are no exceptions nor arbitrary grade adjustments for individual students, nor grade guarantees of any kind, for any reason.

Special Considerations: If a situation exists or arises that you think may hinder your progress in this class, you must notify me as soon as possible.

Advising: For questions related to prerequisites and course suggestions, please contact the Statistics Department's Academic Planning Specialist, Brooke Raake (raake.5@osu.edu).

Academic Misconduct: It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term "academic misconduct includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct http://studentlife.osu.edu/csc/.

Disability Services: Students with disabilities (including mental health, chronic or temporary medical conditions) that have been certified by the Office of Student Life Disability Services will be appropriately accommodated and should inform the instructor as soon as possible of their needs. The Office of Student Life Disability Services is located in 098 Baker Hall, 113 W. 12th Avenue; telephone 614- 292-3307,

slds@osu.edu; slds.osu.edu