Syllabus for MA 588/688
Advanced Statistics – Summer 2020
Instructor: Dr. Shannon Starr
Class Times: MWF 2:40pm – 4:00pm (on Canvas/Zoom)

The topic of this course is statistics from a mathematical perspective. You are presumed to have a background in probability theory or elementary statistics (or both) at the level of an undergraduate or graduate course[1]. Therefore, the course is called advanced statistics.

Statistics is a large subject with many practical technicalities involved. We focus on the mathematical aspects. If your background is in probability theory, then you may think of statistical questions as the inverse problems from probability theory. If your background is in statistics, then you will discover that we focus more on derivations and proofs than in a first elementary statistics course.

All class work will be assigned and collected through Canvas.

Instructor email: slstarr@uab.edu

Office hours: To be announced, held on Zoom.

Textbook: No official textbook is required. We will follow lecture notes written by Dr. Nikolai Chernov, who unfortunately passed away in 2014'.

Dr. Starr, the instructor, will also disseminate additional notes specific to some of the topics for group projects.

Optional additional resources: One textbook that is a good reference if you can get it is

Statistical Inference by George Casella and Roger L. Berger.

A more elementary reference, but which covers some of the same material is

Probability and Statistical Inference by Robert V. Hogg and Elliot A. Tanis,

which comprises the material usually from a first elementary statistics course such as MA 587 or MA 687, which has some overlap with this course.

Description from the schedule of classes: Parameter estimations, maximum likelihood estimation, sufficient statistic, hypothesis testing, Neyman-Pearson Lemma, p-value, Kolmogorov-Smirnov test, Anderson-Darling test, P-P plot, Q-Q plot, testing for distribution type, location and scale parameters, mean squared error.

These topics seem to span approximately Chapters 6 through 12 in Casella and Berger. But we will follow the free online course notes written by Dr. Nikolai Chernov.

Group projects: We will have group projects where you select a topic of current interest in groups of 2–3 students. More details about topics will be listed below. Some of the class time will be used to lecture on topics similar to the group project topic list. One example will be compressed sensing. Another large class of examples come from machine learning. There will also be some instruction on how to make videos, since that is how you will present your project. There will be a short video on a general topic by each group by the middle of the semester. Two weeks before the end of class there will be a paper description of your project due. The final video for the project will be due by the last day of lecture.

The videos and the write-up will be disseminated to other students of the class and some portion of the last HW or the take-home final exam will involve students watching and answering questions about all the other group projects.

[1]If you had MA 485 or MA 585 or MA 587 or MA 687 (or more than one of these) then you automatically qualify for the prerequisites. If you have some other class or combination of classes please do feel free to contact me at my email address and we may discuss whether you should take the class and any background you may want to brush up on.

Homeworks: There will be two types of HW projects: written and computer. The written HW's will focus on more theoretical issues or algebraic derivations. The computerized HW's will focus on numerical types of questions, as well as building computer skills. In the first week of class we will introduce Matlab. All UAB students may download a free copy of Matlab using the university's license. We may also use other online software for computerized HW's if we find that the programs we want are available. All HW's are to be submitted on Canvas (written and computerized). We will discuss how to do that in the first lecture. As a practical matter, HW's will be assigned (and graded) more frequently earlier in the semester. Later in the semester, the work shifts more towards tests and projects.)

Journal: A small percentage of the grade will be reserved for a weekly short summary of what you read in the class notes or additional textbook that week. This will typically be a couple of paragraphs long up to a couple of pages. As with all assignments, the weekly summary will be submitted on Canvas (and will not be shared with the other students of the class).

Participation: Approximately 1 time per week you must either volunteer and answer a question by me during lecture (which I will describe further in class) or participate in a Zoom office hour, including answering some questions by me relating to the last lecture. This will comprise a small percentage of the final grade.

Take home midterm: Approximately half way through the semester we will have a take-home midterm. You must work on the take home exams by yourself, and without consulting any references other than the ones I explicitly state are allowed. As with all other assignments, you will submit your work on Canvas.

Take home final exam: The final exam will be a take-home exam, and it will follow the same guidelines as the midterm. You will be assigned a certain number of problems and you will select a prescribed number of those to answer. (You will have a choice over which problems you want to answer.)

Extra work for MA 688: We will have additional HW problems for MA 688, covering more advanced material. We will also have additional questions on the take-home midterm and take-home final exam. The amount of additional work will be 10-20%.

Grading scheme:

$$\begin{aligned}&(\text{Journal } 5\%) + (\text{Participation } 5\%) + (\text{Computer HW } 10\%) + (\text{Written HW } 20\%)\\ &\qquad + (\text{Midterm } 15\%) + (\text{Final } 20\%) + (\text{Project } 25\%) = (\text{total } 100\%)\end{aligned}$$

List of possible group projects:

① Concentration of measure for nonlinear statistics such as the maximum of numerous linear statistics;

② Statistics for random matrices and their eigenvalues and singular values;

③ q-generalizations of classical statistics distributions such as the beta-distribution;

④ Compressed sensing and the Netflix problem (using concentration of measure);

⑤ Replica symmetry and annealing in a deep Boltzmann machine (following Pierluigi Contucci, et.al.);

⑥ Statistics in modeling (and stochastic differential equations);

⑦ Statistical models of infection spreading (including the contact model and the SIR model).

A longer and more detailed list (with a brief summary for each topic) will be distributed in the first week of class. You may also suggest a topic, and if we discuss it and both agree, then you can do your project on the topic of your own suggestion.