Gaussian and Empirical Process Theory for High-dimensional Data

STAT 591 Special Topics in Statistics, Spring Quarter 2023

Instructor:	Prof. Alexander Giessing giessing@uw.edu	**Office Hours:**	Mon 10:00 AM – 11:00 AM or by appointment (B-308 PDL)
Lectures:	Wed Fri 11:30 – 12:50 PM	**Classroom:**	C-301 PDL
Course Page:	https://canvas.uw.edu/courses/1635483		

Course Description: In this course we develop elements of the theory of Gaussian and empirical processes that have proved useful for statistical inference in high-dimensional models, i.e. statistical models in which the number of parameters is much larger than the sample size. The course consists of three parts, with the first two parts laying the foundation for the third one: an introduction to modern techniques in Gaussian processes, a recap of classical empirical process theory emphasizing weak convergence on metric spaces and $\mathbb{R}^d$, and lastly, a discussion of some modern problems in high-dimensional inference. This course is part of the graduate course offerings at the Pacific Institute for the Mathematical Sciences (PIMS, https://courses.pims.math.ca/tag/2022-2023/).

Prerequisites: STAT 581, 582 (Advanced Theory of Statistical Inference) or MATH 581, 582 (Mathematics of Data Science). A course in measure theoretical probability theory might be helpful.

Lecture Notes: I will post typed lecture notes on Canvas 2-3 days after the lecture. I advise to take notes during the lecture.

References: There is no required textbook, we will use a number of different texts:

- Dudley, R. M. (2014). *Uniform Central Limit Theorems.* Cambridge University Press.
  (General reference.)
- Dudley, R. M. (2002). *Real Analysis and Probability.* Cambridge University Press.
  (General reference, mostly for results that we will not prove.)
- Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models.* Cambridge University Press.
  (Main reference, we will cover material from chapters 2-3.)
- van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes.* Springer.
  (If necessary, we review chapters 2.2–2.6. The plan is to focus on chapters 2.9, 2.10, 2.11, 3.6)
- Nankai Lectures Winter 2018
  (I expect that you are at ease with the material in chapter 2, familiar with the material in chapter 4, and have at least a passing idea what chapter 5 is about.)
- Relevant journal articles for the third part of the lecture will be posted on canvas.

Homework: There will be 4 problem sets. Problem sets will be posted on Canvas. Please submit your solution in a single pdf or jpg. Unexcused late homework submissions will receive a score of zero. You may work in groups of two or three students on the homework problems, please indicate your study group members on your homework submission. However, verbatim copying solutions is strictly forbidden; each student must produce their own solutions.

Final Exam: There will be a final oral exam which works as follows: I will divide the lecture in roughly eight topics which I will share with you ahead of time. At the day of the examination you will randomly draw two topics and give two 10-15 min presentations on these topics on the white board (no prepared notes allowed). Each presentation will conclude with ca. five minutes of follow-up questions.

Schedule and Grade Policy:

Homework (50%) ..... 11:59 PM at due dates
Final (50%) ..... Wed, June 7, 2023

Academic Integrity: Students shall abide by the University of Washington Academic Responsibility policies, which are outlined at https://depts.washington.edu/grading/pdf/AcademicResponsibility.pdf. Violations and suspected violations will be reported to the appropriate Dean's Representative and through the webpage for Community Standards and Student Conduct. The instructor reserves the right to assign a failing grade for the course for serious violations of student conduct.

Academic Accommodations: Your experience in this class is important to me. It is the policy and practice of the University of Washington to create inclusive and accessible learning environments consistent with federal and state law. If you have already established accommodations with Disability Resources for Students (DRS), please activate your accommodations via myDRS so we can discuss how they will be implemented in this course. If you have not yet established services through DRS, but have a temporary health condition or permanent disability that requires accommodations (conditions include but not limited to; mental health, attention-related, learning, vision, hearing, physical or health impacts), contact DRS directly to set up an Access Plan. DRS facilitates the interactive process that establishes reasonable accommodations. Contact DRS at disability.uw.edu.

Religious Accommodations: Washington state law requires that UW develop a policy for the accommodation of student absences or significant hardship due to reasons of faith or conscience or for organized religious activities. The UW's policy, including more information about requesting an accommodation, is available at Religious Accommodations Policy. Accommodations must be requested within the first two weeks of this course using the Religious Accommodations Request form.

Diversity and Inclusion: Diverse backgrounds, embodiments, and experiences are essential to the critical thinking endeavor at the heart of university education. Therefore, I expect you to follow the UW Student Conduct Code in your interactions with your colleagues and me in this course by respecting the many social and cultural differences among us, which may include, but are not limited to: age, cultural background, disability, ethnicity, family status, gender identity and presentation, citizenship and immigration status, national origin, race, religious and political beliefs, sex, sexual orientation, socioeconomic status, and veteran status.

Tentative List of Topics:

1. **Elements of Gaussian processes** (concentration, comparison, anti-concentration, and super-concentration inequalities, Talagrand's Generic chaining bounds). *(ca. 3 weeks)*

2. **Elements of empirical processes** (review of convergence of laws on separable metric spaces, Glivenko-Cantelli and Donsker theorems, permanence theorems, applications to multiplier CLTs and the bootstrap). *(ca. 2 weeks)*

3. **Theoretical problems in high-dimensional inference** (a selection of the following topics: Gaussian approximation, high-dimensional CLTs and multiplier bootstrap when function classes are not Donsker, double machine learning/ debiased inference with high-dim. nuisance parameters, inference on artificial neural nets (ANNs)). *(ca. 5 weeks)*