CSC 696H-001 Topics in Concentration of Measure

Gould-Simpson Rm 701, Monday & Wednesday 12:30pm-1:45pm

Description of Course

This course introduces the concentration of measure phenomenon. Students will learn various techniques to control tail bounds of popular mean estimators and estimators for the minimum value of a function. The course will consist of three part: (i) basic concentration of measure bounds such as subgaussian random variable, sub-exponential random variable, empirical Bernstein, robust mean estimators, and McDiarmid's inequality, (ii) numerically tight confidence bounds including time-uniform bounds, bounds for linear models, PAC-Bayes bounds, and their applications to machine learning problems (iii) advanced tools for the deviation of functions including covering and packing, concentration of Rademacher/Gaussian complexities, metric entropy, and chaining. Throughout, students will not only learn fundamental tools upholding the current understanding of concentration of measure but also develop skills for adapting these techniques to their own research needs such as developing and analyzing new algorithms for interactive machine learning problems like multi-armed bandits and A/B testing.

Course Prerequisites

Students must have strong familiarity with:

- Linear Algebra: linear space, basis, dimensions, linear transformations, matrices, eigenvalues and eigenvectors, positive definiteness of a matrix, matrix decompositions such as the SVD
- Multivariate Calculus: total derivative, gradient, linearity of the derivatives, (second-order) Taylor's expansion
- Basic probability theory: elementary events, definitions of probability, discrete and continuous random variables, distribution laws, (conditional) expectation, (conditional) independence, law of large numbers, central limit theorems.
- Basic programming: fluency in at least one programming language (e.g. Matlab, Julia, Python, C, C++).
- Machine learning theory (recommended but not required): Some concepts are highly related to machine learning theory.

Instructor and Contact Information

Kwang-Sung Jun
Gould-Simpson 746
kjun@cs.arizona.edu
Personal page: https://kwangsungjun.github.io/
Course D2L: https://d2l.arizona.edu/d2l/home/1347795
Course Piazza: https://piazza.com/arizona/fall2023/csc696h001 code: wildcatsxn9b
Gradescope: Entry Code:XVZ56G
Office Hour: Tue 4:30-5:30pm, GS 746.

Obtaining Help

- **Advising:** If you have questions about your academic progress this semester, or your chosen degree program, consider contacting your graduate program coordinator and faculty advisor. Your program coordinator, faculty advisor, and the Graduate Center can guide you toward university resources to help you succeed. **Computer Science students** are encouraged to email gradadvising@cs.arizona.edu for advising related questions.
- **Life challenges:** If you are experiencing unexpected barriers to your success in your courses, please note the Dean of Students Office is a central support resource for all students and may be helpful. The Dean of Students Office can be reached at 520-621-7057 or DOS-deanofstudents@email.arizona.edu.
- **Physical and mental-health challenges**: If you are facing physical or mental health challenges this semester, please note that Campus Health provides quality medical and mental health care. For medical appointments, call 520-621-9202. For After Hours care, call (520) 570-7898. For the Counseling & Psych Services (CAPS) 24/7 hotline, call (520) 621-3334.
- **UA Ombuds**: The UA Ombuds Office (https://ombuds.arizona.edu/) helps with a wide variety of issues, concerns, questions, conflicts, and challenges. The primary mission of the Ombuds Program is to assist individuals in resolving conflict, facilitating communication, and assisting the University by surfacing issues and providing feedback on emerging or systemic concerns. Communications with the Ombuds Committee are informal and off-the-record. The Ombuds Committee is governed by the following standards: (1) Confidentiality; (2) Impartiality: (3) Informality; and (4) Independence.

Class Recordings

For lecture recordings, which are used at the discretion of the instructor, students must access content in D2L only. Students may not modify content or re-use content for any purpose other than personal educational reasons. All recordings are subject to government and university regulations. Therefore, students accessing unauthorized recordings or using them in a manner inconsistent with UArizona values and educational policies (Code of Academic Integrity and the Student Code of Conduct) are also subject to civil action.

The class will be recorded but **will not be livestreamed** since the course is designated as being in-person. The recording will be available after the class. Students are highly recommended to attend the class in-person to maximize the learning outcome.

If students do not wish to be identified by name, please contact the instructor directly.

Course Objectives

A successful student will be able to explain the key concepts, strengths, and weaknesses of various topics in concentration of measure:

- Sub-Gaussian random variables
- Sub-exponential random variables
- Robust mean estimators
- Empirical Bernstein inequality
- Martingale-based methods
- Lipschitz functions of Gaussian random variables

- Time-uniform bounds and applications to pure exploration
- Prediction error bounds for linear models and linear bandits
- Numerically tight confidence bounds (e.g., KL-divergence-based or betting-based bounds)
- PAC-Bayes
- PAC-Bayes applied to offline evaluation of recommendation systems.
- Uniform laws of large numbers
- Upper bounds on the Rademacher complexity
- Covering and packing
- Rademacher/Gaussian complexity
- Metric entropy
- Sub-Gaussian processes
- Chaining

For a more fine-grained description of the learning objectives, see the week-by-week schedule below. This course covers key topics that are not only useful for estimating means but also relevant in modern machine learning research. The students can directly adapt learned topics to their own research. Students are expected to dedicate a significant amount of time on understanding mathematical concepts and skills outside the classroom.

Expected Learning Outcomes

The expected learning outcomes of the course are:

- To explain what is sub-Gaussianity and how it is compared to Gaussian random variables.
- To explain how the sub-Gaussian and the sub-exponential concentration are different and why the extra lower order term in the latter is necessary in the worst case.
- To explain why robust mean estimators can have sub-Gaussian tails.
- To explain the strengths and weaknesses of McDiarmid's inequality compared to the sub-Gaussian inequality.
- To relate KL divergence based bounds to betting-based bounds.
- To explain how a union bound over a finite number of members can be related to PAC-Bayes bounds.
- To explain covering and packing.
- To explain chaining.

Absence and Class Participation Policy

The UA's policy concerning Class Attendance, Participation, and Administrative Drops is available at https://catalog.arizona.edu/policy/class-attendance-and-participation

The UA policy regarding absences for any sincerely held religious belief, observance or practice will be accommodated where reasonable: http://policy.arizona.edu/human-resources/religious-accommodation-policy.

Absences pre-approved by the UA Dean of Students (or dean's designee) will be honored. See https://deanofstudents.arizona.edu/policies/attendance-policies-and-practices

Participating in the course and attending lectures and other course events are vital to the learning process. As such, attendance is required at all lectures. Absences may affect a student's final course grade. If you anticipate being absent, are unexpectedly absent, or are unable to participate in class online activities, please contact me as soon as possible. To request a disability-related accommodation to this attendance policy, please contact the Disability Resource Center at (520) 621-3268 or drc-info@email.arizona.edu. If you are experiencing unexpected barriers to your success in your courses, the Dean of Students Office is a central support resource for all students

and may be helpful. The Dean of Students Office is located in the Robert L. Nugent Building, room 100, or call 520-621-7057.

The instructor reserves the right to check attendance in any lecture, which may factor in the participation scores.

Illnesses and Emergencies

- If you feel sick, or may have been in contact with someone who is infectious, stay home. Except for seeking medical care, avoid contact with others and do not travel.
- Notify your instructor(s) if you will be missing up to one week of course meetings and/or assignment deadlines.
- If you must miss the equivalent of more than one week of class and have an emergency, the Dean of Students is the proper office to contact (DOS-deanofstudents@email.arizona.edu). The Dean of Students considers the following as qualified emergencies: the birth of a child, mental health hospitalization, domestic violence matter, house fire, hospitalization for physical health (concussion/emergency surgery/coma/COVID-19 complications/ICU), death of immediate family, Title IX matters, etc.
- Please understand that there is no guarantee of an extension when you are absent from class and/or miss a deadline.

Makeup Policy for Students Who Register Late

If you register late for this class, contact me as soon as you do. You will be expected to submit all missed homework assignments within a week of your registration. It is your responsibility to catch up to the class content.

Course Communications

We will use D2L for communications, but all the discussions will be done in Piazza.

Required Texts and Materials

This includes books, lab materials, field trips, tickets, technology, or any associated costs that are required for students in your class. It is recommended that you also provide the estimated cost of these items for the class, based on pricing through the UA Bookstore.

Textbooks:

- [W] Martin J. Wainwright, High-Dimensional Statistics: A Non-Asymptotic Viewpoint, 2019.
- [BLM] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. Concentration inequalities: A nonasymptotic theory of independence, 2012.

Both are available through UA library. The main textbook will be [W], but we will occasionally use [BLM] and other materials. The instructor will announce other related materials in D2L.

Scheduled Topics/Activities

List topics in logical units in a weekly/daily schedule, including assignment due dates and exam dates.

Instructors must include:

1) *the handout and turn-in dates for all programming assignments, quizzes, etc;*
2) *relevant readings for each lecture/week/topic (readings can be from an assigned textbook, online resources, lecture slides, research paper, etc. - as long as the student has a written resource to turn to should they have problems understanding the lecture)*

Week	Date	#	Description	Additional readings
1	08/21	1	Lecture: Introduction, motivation, course mechanics	
	08/23	2	Lecture: Overview of part 1: basic concentration of measure results. Sub-Gaussian random variables	[W] 2.1.1-2.1.2
2	08/28	3	Lecture: Sub-exponential random variables (1)	[W] 2.1.3-2.1.4
	08/30	4	Lecture: Sub-exponential random variables (2)	[W] 2.1.3-2.1.4
3	09/04		No lecture; labor day.	
	09/06	5	Lecture: Robust mean estimators	Catoni, "Challenging the empirical mean and empirical variance: A deviation study", 2012.
4	09/11	6	Lecture: Empirical Bernstein inequality **Homework 1 assigned**	Maurer, Pontil, " Empirical Bernstein Bounds and Sample Variance Penalization", 2009.
	09/13	7	Lecture: Martingale-based methods and McDiarmid's inequality.	[W] 2.2
5	09/18	8	Lecture: Lipschitz functions of Gaussian variables.	[W] 2.3
	09/20	9	Lecture: Overview of part 2: numerically tight concentration of measure results. Time-uniform bounds and applications in multi-armed bandits.	Howard et al. "Time-Uniform, Nonparametric, Nonasymptotic Confidence Sequences", 2021.
6	09/25	10	Lecture: Bounds for linear models and linear bandits (1) **Homework 1 due**	Abbasi-Yadkori et al. "Improved Algorithms for Linear Stochastic Bandits", 2011.
	09/27	11	Lecture: Bounds for linear models and linear bandits (2)	
7	10/02	12	Lecture: Numerically tight bounds: KL-divergence and betting-based bounds.	Cappe et al. "Kullback–Leibler Upper Confidence Bounds For Optimal Sequential Allocation", 2013.
	10/04	13	Lecture: PAC-Bayes bounds (1)	Alquier, "User-friendly introduction to PAC-Bayes bounds", 2021.
8	10/09	14	Lecture: PAC-Bayes bounds (2) with the application to offline evaluation of recommendation systems. **Homework 2 assigned**	London, Sandler, "Bayesian Counterfactual Risk Minimization", 2019
	10/11	15	Lecture: Overview of part 3: advanced tools for generalization bounds. Uniform law of large numbers.	[W] 4.1-4.2

9	10/16	16	Lecture: Upper bounds on the Rademacher complexity **Presentation proposal due**	[W] 4.3
	10/18	17	Lecture: Covering and Packing	[W] 5.1
10	10/23	18	Lecture: Gaussian and Rademacher complexity **Homework 2 due**	[W] 5.2
	10/25	19	Lecture: Metric entropy and sub-Gaussian processes	[W] 5.3-5.4
11	10/30	20	Lecture: Sudakov's lower bound	[W] 5.5
	11/01	21	Lecture: Chaining	[W] 5.6
12	11/06	22	Student paper presentation	
	11/08	23	Student paper presentation	
13	11/13	24	Student paper presentation	
	11/15	25	Student paper presentation	
14	11/20	26	Student paper presentation	
	11/22	27	Student paper presentation	
15	11/27	28	Student paper presentation	
	11/29	29	Student paper presentation	
16	12/04	30	Student paper presentation	
	12/06	31	Student paper presentation	

Final Examination or Project

The date and time of the final exam or project, along with links to the Final Exam Regulations and Final Exam Schedule: https://registrar.arizona.edu/faculty-staff-resources/room-class-scheduling/schedule-classes/final-exams

The final project is the student paper presentation. This must be done individually. The student must consult the instructor to choose a paper on concentration inequalities. The presentation will be 50 minutes followed by questions and answers. The presentation must include a clear exposition of the background, the problem being addressed, the solution the paper proposes, a comparison of the result with prior arts, key technical details/proofs, and possible extensions and open problems. There will be an extra 10 minutes for addressing the questions and weaknesses mentioned by the paper critiques (see grading scale and policies for more on the paper critique).

The final project will be evaluated based on

- The quality of the background (motivation, application, etc.) being presented.
- The quality of the problem definition being presented.
- The quality of the comparisons with the state-of-the-art results.
- The quality of the technical depth (e.g., proof) being presented.
- How the

Grading Scale and Policies

Specify the grade distribution for the course. University policy regarding grades and grading systems is available at http://catalog.arizona.edu/policy/grades-and-grading-system

Provide a detailed explanation of the methods of evaluation and how the final grade will be

calculated, including components/assignments, weightings, evaluation criteria, explanation of how late work will be graded, and description of extra-credit opportunities.

The instructing staff will grade your homework assignments and project on a scale from 0 to 100, with the following weights:

- Homework assignments: 40% (20% each)
- Participation: 5%
- Weekly discussion: 5%
- Quiz: 10%
- Paper critique: 10%
- Presentation: 30%
    - 10% proposal
    - 20% presentation evaluation

The final grade in the course is determined by the better of a per-class grading curve and overall performance:

- 90% or better: A;
- 80% or better: B;
- 70% or better: C;
- 60% or better: D;
- below 60%: E.

See the "scheduled topics / activities" section for the timeline and due dates of the homework assignments and the project. Every homework is due in 2 weeks, and will be returned to students within 7 week days except for the university holidays and before the next homework is due. Grading delays beyond promised return-by dates will be announced as soon as possible with an explanation for the delay.

There will be **no late days** for homework assignments. No exceptions will be made except for extraordinary situations.

There will be a popup quiz every 1-2 weeks.

Participation score will be based on the attendance and the Piazza activities: asking questions, answering them, and posting discussions.

Presentation proposal will be a one-page summary of why the student is interested in presenting the paper and the preliminary plans on what to focus on for the presentation. Again, there will be no late days.

Once the presentation topics are decided, for each presentation topic, two other students will be selected at random. These two students will have to submit an independent paper critique (2-3 pages), which will take the standard conference paper review format such as NeurIPS/ICML. The deadline is 12:30pm before the presentation day (For Monday presentations, the deadline is Friday 12:30pm).

Grades for the presentation proposal and the presentation will be available in 7 weekdays. The final grade will be posted within 48 hours of the last class.

Incomplete (I) or Withdrawal (W):

Requests for incomplete (I) or withdrawal (W) must be made in accordance with University policies, which are available at http://catalog.arizona.edu/policy/grades-and-grading-system#incomplete and http://catalog.arizona.edu/policy/grades-and-grading-system#Withdrawal respectively.

Dispute of Grade Policy: If you wish to dispute your grade for homework assignments or the presentation, you have two weeks after the grade has been turned in. In addition, even if only the student disputes one portion of the grading for that unit, the instructor reserves the right to revisit the entire unit (homework assignment or project).

Department of Computer Science Code of Conduct

The Department of Computer Science is committed to providing and maintaining a supportive educational environment for all. We strive to be welcoming and inclusive, respect privacy and confidentiality, behave respectfully and courteously, and practice intellectual honesty. Disruptive behaviors (such as physical or emotional harassment, dismissive attitudes, and abuse of department resources) will not be tolerated. The complete Code of Conduct is available on our department web site. We expect that you will adhere to this code, as well as the UA Student Code of Conduct, while you are a member of this class.

Classroom Behavior Policy

To foster a positive learning environment, students and instructors have a shared responsibility. We want a safe, welcoming, and inclusive environment where all of us feel comfortable with each other and where we can challenge ourselves to succeed. To that end, our focus is on the tasks at hand and not on extraneous activities (e.g., texting, chatting, reading a newspaper, making phone calls, web surfing, etc.).

Students are asked to refrain from disruptive conversations with people sitting around them during lecture. Students observed engaging in disruptive activity will be asked to cease this behavior. Those who continue to disrupt the class will be asked to leave lecture or discussion and may be reported to the Dean of Students.

Threatening Behavior Policy

The UA Threatening Behavior by Students Policy prohibits threats of physical harm to any member of the University community, including to oneself. See http://policy.arizona.edu/education-and-student-affairs/threatening-behavior-students.

Accessibility and Accommodations

At the University of Arizona, we strive to make learning experiences as accessible as possible. If you anticipate or experience barriers based on disability or pregnancy, please contact the Disability Resource Center (520-621-3268, https://drc.arizona.edu/) to establish reasonable accommodations.

Code of Academic Integrity

Students are encouraged to share intellectual views and discuss freely the principles and applications of course materials. However, graded work/exercises must be the product of independent effort unless otherwise instructed. Students are expected to adhere to the UA Code of Academic Integrity as described in the UA General Catalog. See https://deanofstudents.arizona.edu/student-rights-responsibilities/academic-integrity .

Uploading material from this course to a website other than D2L (or the class piazza) is strictly prohibited and will be considered a violation of the course policy and a violation of the code of academic integrity. Obtaining material associated with this course (or previous offerings of this course) on a site other than D2L (or the class piazza), such as Chegg, Course Hero, etc. or accessing these sites during a quiz or exam is a violation of the code of academic integrity. Any student

determined to have uploaded or accessed material in an unauthorized manner will be reported to the Dean of Students for a Code of Academic Integrity violation, with a recommended sanction of a failing grade in the course.

The University Libraries have some excellent tips for avoiding plagiarism, available at https://new.library.arizona.edu/research/citing/plagiarism.

Selling class notes and/or other course materials to other students or to a third party for resale is not permitted without the instructor's express written consent. Violations to this and other course rules are subject to the Code of Academic Integrity and may result in course sanctions. Additionally, students who use D2L or UA e-mail to sell or buy these copyrighted materials are subject to Code of Conduct Violations for misuse of student e-mail addresses. This conduct may also constitute copyright infringement.

This course has **zero tolerance policy** on the academic integrity issue – any form of academic integrity that was not discussed explicitly with the instructor prior to the incident will result in an F grade. If you have any doubt, please discuss it with the instructor first before taking actions.

Nondiscrimination and Anti-harassment Policy

The University of Arizona is committed to creating and maintaining an environment free of discrimination. In support of this commitment, the University prohibits discrimination, including harassment and retaliation, based on a protected classification, including race, color, religion, sex, national origin, age, disability, veteran status, sexual orientation, gender identity, or genetic information. For more information, including how to report a concern, please see http://policy.arizona.edu/human-resources/nondiscrimination-and-anti-harassment-policy

Our classroom is a place where everyone is encouraged to express well-formed opinions and their reasons for those opinions. We also want to create a tolerant and open environment where such opinions can be expressed without resorting to bullying or discrimination of others.

Additional Resources for Students

UA Academic policies and procedures are available at http://catalog.arizona.edu/policies

Visit the UArizona COVID-19 page for regular updates.

Campus Health

http://www.health.arizona.edu/

Campus Health provides quality medical and mental health care services through virtual and in-person care. Voluntary, free, and convenient COVID-19 testing is available for students on Main Campus. COVID-19 vaccine is available for all students at Campus Health.

Phone: 520-621-9202

Counseling and Psych Services (CAPS)

https://health.arizona.edu/counseling-psych-services

CAPS provides mental health care, including short-term counseling services.

Phone: 520-621-3334

The Dean of Students Office's Student Assistance Program

https://deanofstudents.arizona.edu/support/student-assistance

Student Assistance helps students manage crises, life traumas, and other barriers that impede success. The staff addresses the needs of students who experience issues related to social adjustment, academic challenges, psychological health, physical health, victimization, and relationship issues, through a variety of interventions, referrals, and follow up services.

Email: DOS-deanofstudents@email.arizona.edu

Phone: 520-621-7057

Survivor Advocacy Program
https://survivoradvocacy.arizona.edu/
The Survivor Advocacy Program provides confidential support and advocacy services to student survivors of sexual and gender-based violence. The Program can also advise students about relevant non-UA resources available within the local community for support.
Email: survivoradvocacy@email.arizona.edu
Phone: 520-621-5767

Safety on Campus and in the Classroom (required for all in-person instruction)

For a list of emergency procedures for all types of incidents, please visit the website of the Critical Incident Response Team (CIRT): https://cirt.arizona.edu/case-emergency/overview

Also watch the video available at
https://arizona.sabacloud.com/Saba/Web_spf/NA7P1PRD161/common/learningeventdetail/crtfy000000000003560

Confidentiality of Student Records

http://www.registrar.arizona.edu/ferpa

Land Acknowledgement Statement

We respectfully acknowledge the University of Arizona is on the land and territories of Indigenous peoples. Today, Arizona is home to 22 federally recognized tribes, with Tucson being home to the O'odham and the Yaqui. Committed to diversity and inclusion, the University strives to build sustainable relationships with sovereign Native Nations and Indigenous communities through education offerings, partnerships, and community service.

Subject to Change Statement

Information contained in the course syllabus, other than the grade and absence policy, may be subject to change with advance notice, as deemed appropriate by the instructor.