CREST - GENES
Cours doctoraux 2022 – 2023

Statistics of Differential Privacy

Lukas Steinberger
University of Vienna

SCHEDULE	Monday	7th November 2022 14th November 2022	From 9:00 to 12:15	Room 2045 Room 2045
	Thursday	10th November 2022 17th November 2022	From 10:00 to 12:15	Room 2045 Room 2033

Summary:

With the advent of big data and powerful machine learning methods, traditional approaches towards data privacy protection, such as aggregation and anonymization, have reached their limits. At the turn of the century a new notion of data privacy protection called `differential privacy' (DP) has emerged that is inherently statistical and admits a rigorous mathematical definition and analysis. Despite its increasing popularity in recent years, statisticians are only beginning to explore its full potential and its limitations for statistical inference. In this course we give an introduction to the rapidly growing field of differential privacy from the perspective of mathematical statistics and formally study trade-offs between privacy protection and data utility for several parametric and nonparametric estimation problems. After clarifying basic concepts and looking at first examples of differentially private data release mechanisms, we will mainly focus on the mathematical techniques required to establish statistical optimality properties. We will also encounter several open problems and possible directions for future research. In particular, we will cover the following topics:

- General introduction and properties of DP data release mechanisms
- Basic principles of mechanism design
- Central vs. local paradigm of DP
- Strong data processing inequalities
- Theory of minimaxity with local DP
- Towards a theory of efficiency with local DP (time permitting)

The prerequisites for this course are real analysis and probability theory. Familiarity with basic concepts of mathematical statistics will be very helpful but not strictly necessary.

References:

- **Dwork, C., McSherry, F., Nissim, K. and Smith, A. (2006).** Calibrating noise to sensitivity in private data analysis. In Theory of Cryptography (S. Halevi and T. Rabin, eds.). Lecture Notes in Computer Science 265–284. Springer.

- **Wasserman, L., and Zhou, S. (2010).** A statistical framework for differential privacy. *J. Amer. Statist. Assoc.* 105, 375-389.
- **Duchi, J. C., Jordan, M. I. and Wainwright, M. J. (2018).** Minimax optimal procedures for locally private estimation. J. Amer. Statist. Assoc. 113, 182–201.
- **Rohde, A. and Steinberger, L. (2020).** Geometrizing rates of convergence under local differential privacy constraints. *Annals of Statistics* 48(5)*, 2646–2670.*
- **Butucea, C., Rohde, A. and Steinberger, L. (2022).** Interactive versus non-interactive locally differentially private estimation: Two elbows for the quadratic functional. *arXiv:2003.04773*

Evaluation:

Participants who take the course for credits are expected to propose a small project where they apply some of the methods learned during the course.

Ces cours sont proposés aux étudiants de 3$^{\text{ème}}$ année de l'ENSAE, de l'ENSAI, ouverts aux étudiants de M2 ou inscrits en *thèse*. ***Une inscription préalable est demandée impérativement*** *pour tous les étudiants de l'ENSAE, de l'ENSAI, ou* extérieurs, à Lyza RACON : lyza.racon@ensae.fr ou par téléphone **au 0170266926** **afin de pouvoir être admis dans les locaux de l'ENSAE** **et pouvoir être joints en cas de nécessité par les organisateurs du cours.**