

Title: Fairness in machine learning: a study of the Demographic Parity constraint

Abstract: In various domains, statistical algorithms trained on personal data take pivotal decisions which influence our lives on a daily basis. Recent studies show that a naive use of these algorithms in sensitive domains may lead to unfair and discriminating decisions, often inheriting or even amplifying biases present in data. In the first part of the talk, I will introduce and discuss the question of fairness in machine learning through concrete examples of biases coming from the data and/or from the algorithms. In a second part, I will demonstrate how statistical learning theory can help us better understand and overcome some of those biases. In particular, I will present a selection of recent results from two of my papers on the Demographic Parity constraint, a popular fairness constraint. In particular I will describe an interesting link between this constraint and optimal transport.

References:

- A minimax framework for quantifying risk-fairness trade-off in regression (with E. Chzhen), *Ann. Statist.* 50(4): 2416-2442 (Aug. 2022). DOI: 10.1214/22-AOS2198;
- Fair learning with Wasserstein barycenters for non-decomposable performance measures (with S. Gaucher and E. Chzhen), *AISTATS* 2023.