

Stanford University
Department of Statistics

DEPARTMENTAL SEMINAR

4:30pm, Tuesday, May 7, 2019

Herrin Hall Room T175

Refreshments served at 4pm in Sequoia Lounge.

Speaker: Caroline Uhler, *MIT*

Title: From Causal Inference to Autoencoders and Gene Regulation

Abstract:

Recent progress in genomics makes it possible to perform perturbation experiments at a very large scale. This motivates the development of a causal inference framework that is based on observational and interventional data. We characterize the causal relationships that are identifiable and present the first provably consistent algorithm for learning a causal network from such data. I will then couple gene expression with the 3D genome organization. In particular, we will discuss approaches for integrating different data modalities such as sequencing or imaging via autoencoders. We end by a theoretical analysis of autoencoders linking overparameterization to memorization. In particular, we will show that overparameterized single-layer fully connected autoencoders as well as deep convolutional autoencoders memorize images, i.e., they produce outputs in the span of the training images. Collectively, this talk will highlight the symbiosis between genomics and machine learning, showing how biology can lead to new theorems, which in turn can guide biological experiments.