Stanford University
Departments of Mathematics and Statistics

PROBABILITY SEMINAR

4pm, Monday, April 9, 2018
Sequoia Hall Room 200
Refreshments served at 3:30pm in the Lounge.

Speaker: Andrea Montanari  
Department of Statistics and  
Department of Electrical Engineering,  
Stanford University

Title: A Mean Field View of the Landscape of  
Two-Layers Neural Networks

Abstract:

Multi-layer neural networks are among the most powerful models in machine learning and yet, the fundamental reasons for this success defy mathematical understanding. Learning a neural network requires to optimize a highly non-convex and high-dimensional objective (risk function), a problem which is usually attacked using stochastic gradient descent (SGD). In this talk, I will discuss how some tools from probability theory and analysis can be useful to tackle for this challenge. I consider a simple case, namely two-layers neural networks, and prove that — in a suitable scaling limit— SGD dynamics is captured by a certain non-linear partial differential equation. This PDE can be viewed as a gradient flow in a space of probability measures.

I will then study a few simple examples, and show how the asymptotic description can be used to prove convergence of SGD to network with nearly-ideal generalization error. This description allows to “average-out” some of the complexities of the landscape of neural networks, and can be used to describe noisy SGD as well.

This is based on joint work with Song Mei and Phan-Minh Nguyen.