

Estimating Probability Distributions and their Properties

Abstract:

This thesis studies several theoretical problems in nonparametric statistics and machine learning, mostly in the areas of density functional estimation (estimating an integral functional of the population distribution from which the data are drawn) and density estimation (estimating the entire population distribution from which the data are drawn). A consistent theme is that, although nonparametric density estimation is traditionally thought to be intractable in high-dimensions, several equally useful tasks are relatively more tractable, even with similar or weaker assumptions on the distribution.

My work on density functional estimation focuses on several types of integral functionals, such as information-theoretic quantities (entropies, mutual informations, and divergences), measures of smoothness, and measures of distance between distributions, which play important roles as subroutines elsewhere in statistics, machine learning, and signal processing. For each of these quantities, under a variety of nonparametric models, I provide some combination of (a) new estimators, (b) upper bounds on convergence rates of these new estimators, (c) new upper bounds on the convergence rates of established estimators, (d) concentration bounds or asymptotic distributions for estimators, or (e) lower bounds on the minimax risk of estimation.

For density estimation, whereas the majority of prior work has focused on estimation under L^p losses, I consider minimax convergence rates under several new losses, including Wasserstein distances and a large class of metrics called integral probability metrics (IPMs) that includes, for example, L^p , total variation, Kolmogorov-Smirnov, earth-mover, Sobolev, Besov, and some RKHS distances. These losses open several new possibilities for nonparametric density estimation in certain cases, including, for example, convergence rates with no or reduced dependence on dimension. The main results here are the derivation of minimax convergence rates, but I also explore several consequences. For example, I show that IPMs have close connections with generative adversarial networks (GANs) and leverage my results to prove the first finite-sample guarantees for GANs, showing minimax optimality in an idealized model of GANs as density estimators. These results may help explain why these tools appear to perform well at problems that are intractable from traditional perspectives of nonparametric statistics.



Speaker:

Shashank Singh

Thesis Committee:

Barnabas Poczos, Chair

Ryan Tibshirani

Larry Wasserman

Bharath Sriperumbudur (Pennsylvania State University)

Aug. 15, 2019

12:00pm

GHC 6115

Link to draft document:

http://www.contrib.andrew.cmu.edu/~sss1/publications/singh_thesis.pdf

