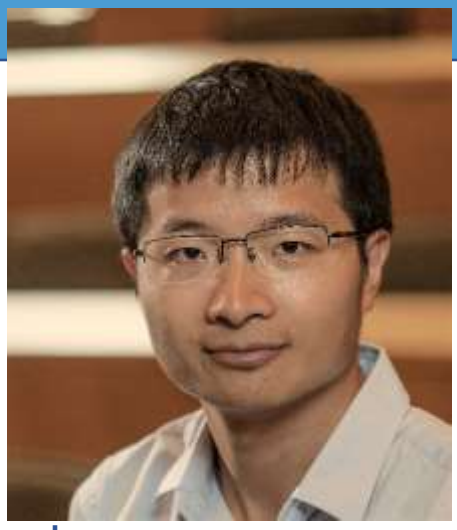
BIOSTATISTICS SEMINAR

Gaussian Differential Privacy



Weijie Su, PhD

Wharton Statistics and Data Science Department

University of Pennsylvania

Privacy-preserving data analysis has been put on a firm mathematical foundation since the introduction of differential privacy (DP) in 2006. This privacy definition, however, has some well-known weaknesses: notably, it does not tightly handle composition. In this talk, we propose a relaxation of DP that we term "f-DP", which has a number of appealing properties and avoids some of the difficulties associated with prior relaxations. This relaxation allows for lossless reasoning about composition and post-processing, and notably, a direct way to analyze privacy amplification by subsampling. We define a canonical single-parameter family of definitions within our class that is termed "Gaussian Differential Privacy", based on hypothesis testing of two shifted normal distributions. We prove that this family is focal to f-DP by introducing a central limit theorem, which shows that the privacy guarantees of any hypothesis-testing based definition of privacy converge to Gaussian differential privacy in the limit under composition. We also demonstrate a central limit theorem phenomenon for high-dimensional query answering, which gives rise to an uncertainty principle style result showing that for any mechanism, the product of its privacy guarantee and estimation loss is lower bounded by the dimension. Finally, we demonstrate the use of the tools we develop by giving an improved analysis of the privacy guarantees of noisy stochastic gradient descent. This is based on joint work with Jinshuo Dong, Aaron Roth, Zhiqi Bu, Qi Long, and Linjun Zhang.

Thursday February 3, 2022, 3:30-4:30 PM Eastern – Virtual using link and info below.

Link: https://unc.zoom.us/j/98412143955?pwd=a1p6c3hvZ28wSnk3dVVXQWl0dEpzdz09

Meeting ID: 984 1214 3955 Passcode: 0375501630