Simulation-based Differentially Private Inference for Categorical Data

A seminar by Roberto Molinari

Department of Mathematics and Statistics - Auburn University (USA)

Thursday 27 Jun 2024 | 2.30 p.m.
Room Benvenuti
Department of Statistical Sciences

Differential privacy (DP) provides a mathematical framework for defining a provable disclosure risk in the presence of arbitrary adversaries: it guarantees that whether an individual is in a database or not, the results of a DP procedure should be similar in terms of their probability distribution. While DP mechanisms are provably effective in protecting privacy, they often negatively impact the precision of the statistics computed from them as well as the possibility of performing reliable inferences on them. To address this problem, ideas from simulation-based methods (such as indirect inference) are investigated to deliver easily computable and reliable inference quantities for different statistical tasks. The preliminary numerical and theoretical results are described when employing these approaches for inference on proportions, starting from the standard one-sample proportion test for which only a few solutions exist in the DP framework. These results are also discussed for the two-sample proportion test for which, to the best knowledge, no solution currently exists under a DP setting. Highlighting the good properties of these solutions in terms of the level and power of the tests, it also discusses how these results could be extended to logistic regression with categorical predictors. These results motivate the current work which is being made in this direction.

Università degli Studi di Padova

Dipartimento di Scienze Statistiche