University of California, Irvine
Statistics Seminar

Goodness of Fit Tests for Differentially-Private Frequency Tables

Naisyin Wang
Professor
Department of Statistics
University of Michigan

4 p.m., Thursday, Oct. 20, 2022
6011 Donald Bren Hall

When releasing data to the public, a vital concern is the risk of exposing personal information of the individuals who have contributed to the data set. Many mechanisms have been proposed to protect individual privacy, though less attention has been dedicated to practically conducting valid inferences on the altered privacy-protected data sets. For frequency tables, the privacy-protection-oriented perturbations often lead to negative cell counts. Releasing such tables can undermine users' confidence in the usefulness of such data sets. We focus on releasing one-way frequency tables. We consider a mechanism that satisfies epsilon-differential privacy (DP) without suffering from having negative cell counts. The procedure is optimal in the sense that the expected utility is maximized under a given privacy constraint. Valid inference procedures for testing goodness-of-fit are developed for this and other additive DP privacy-protected procedures. In particular, we propose a de-biased test statistic and derive its asymptotic distribution. We further consider common users' practices such as merging related or neighboring cells or integrating statistical information obtained across different data sources and derive valid testing procedures when these operations occur. Simulation studies show that our inference results hold well even when the sample size is relatively small. Comparisons with the current field standards, including the Laplace, the Gaussian (both with/without post-processing of replacing negative cell counts with zeros), and the Binomial-Beta McClure-Reiter mechanisms, are carried out. We apply the DP methods to the National Center for Early Development and Learning's (NCEDL) multi-state studies data to demonstrate the practical applicability. The work is jointly done with Chengcheng Li and Gongjun Xu.