

1057-62-412

Slavkovic B Aleksandra* (sesa@stat.psu.edu) and **Karwa Vishesh** (vishesh@psu.edu).

Algebraic statistics framework for causal inference and data privacy with discrete data. Preliminary report.

We present an algebraic computational framework that handles special cases of latent class analyses. Specifically, we consider discrete data problems with unobserved variables such that arbitrary linear constraints are imposed on the possible realizations of the complete data, and thus on the possible states of the joint distribution of all the variables (observed and unobserved) in the analysis. The constraints are imposed either by the modeling assumptions, the structure of the latent variables or for consistency reasons. We illustrate our methods by applying them to two important related problems. The first problem pertains to the assessment of disclosure risk of releasing potentially sensitive information from a latent class analysis in the form of class membership probabilities and probability distribution of covariates conditional on the classes. The second problem pertains to estimation of average causal effect in presence of unobserved confounders, under the Neyman-Rubin framework of potential outcomes. The code is implemented in R, but interfaces with 4ti2. (Received January 26, 2010)