

Semiparametric Double Robust M-Estimation in Count Regression with Excess Zeros and Missing Data

Martin Lukusa (Ma-Ding)
Department of Statistics, NCKU

摘要

Typically, count data are likely to exhibit an excess of zeros in various fields including species abundance, epidemiology, social science, traffic accidents, manufacturing, and so on. We consider the problem of estimation for count regression with excess of zeros when the missingness process is missing at random (MAR). In this case, the naive estimator is biased. Although the inverse probability weighting estimation yields good results, but it may suffer from loss of efficiency and less robustness. In this study, we propose four weighting estimators of zero-inflated (excess zeros) Poisson regression. Each proposed estimator solves a particular equation which requires two nonparametric plug-in nuisance components including the selection probability (weight) and the augmentation term. Specifically, the Nadaraya kernel and the generalized additive models are the nonparametric models used to model the nuisance components. The resulting weighting M-estimators are semiparametric, double robust, consistent and relatively more efficient than their parametric counterpart and their related inverse probability weighting estimators. We scrutinized the performance of the proposed estimators in finite sample by conducting a simulation study and exploring two road traffic data sets. Overall, the performance of the proposed methodology was satisfactory.

Keywords: Count data, Excess zeros, Semiparametric estimator, Missing at random, Non-parametric nuisance function, Estimating equation.