Semiparametric modelling of two-component mixtures with stochastic dominance

Jingjing Wu[1] · Tasnima Abedin[2] · Qiang Zhao[3]

Abstract

In this work, we studied a two-component mixture model with stochastic dominance constraint, a model arising naturally from many genetic studies. To model the stochastic dominance, we proposed a semiparametric modelling of the log of density ratio. More specifically, when the log of the ratio of two component densities is in a linear regression form, the stochastic dominance is immediately satisfied. For the resulting semiparametric mixture model, we proposed two estimators, maximum empirical likelihood estimator (MELE) and minimum Hellinger distance estimator (MHDE), and investigated their asymptotic properties such as consistency and normality. In addition, to test the validity of the proposed semiparametric model, we developed Kolmogorov–Smirnov type tests based on the two estimators. The finite-sample performance, in terms of both efficiency and robustness, of the two estimators and the tests were examined and compared via both thorough Monte Carlo simulation studies and real data analysis.

✉ Jingjing Wu
jinwu@ucalgary.ca

1   Department of Mathematics and Statistics, University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada

2   Clinical Research Unit and Translational Laboratories, Alberta Health Services, 1331 29 Street NW, Calgary, AB T2N 4N2, Canada

3   School of Mathematics and Statistics, Shandong Normal University, No.1 University Road, Science Park, Changqing District, Jinan 250358, Shandong, China