Title: Semiparametric imputation using conditional Gaussian mixture models under item nonresponse

Authors: Danhyang Lee (Presenter) and Jae Kwang Kim

Abstract

Imputation is a popular technique for handling item nonresponse. Parametric imputation is based on a parametric model for imputation and is not robust against the failure of the imputation model. Nonparametric imputation is fully robust but is not applicable when the dimension of covariates is large due to the curse of dimensionality. Semiparametric imputation is another robust imputation based on a flexible model where the number of model parameters can increase with the sample size. In this paper, we propose a new semiparametric imputation based on a more flexible model assumption than the Gaussian mixture model. In the proposed mixture model, we assume a conditional Gaussian model for the study variable given the auxiliary variables, but the marginal distribution of the auxiliary variables is not necessarily Gaussian. The proposed mixture model is more flexible and achieves a better approximation than the Gaussian mixture models. The proposed method is applicable to high dimensional covariate problem by including a penalty function in the conditional log-likelihood function. The proposed method is applied to the 2017 Korean Household Income and Expenditure Survey conducted by Statistics Korea.

Keywords: density ratio model, Kullback-Leibler divergence, survey sampling