Small area estimation for grouped data

Yuki Kawakubo*and Genya Kobayashi*

Abstract

We are concerned with small area estimation for grouped data or frequency distribution. There are two fundamental models for model based small area estimation: one is the Fay–Herriot model for area level data, the other is the nested error regression model for unit level data. Because it is difficult to access unit level data in many cases, Fay–Herriot model is more widely used in practice. However, it is often the case that we can observe the frequency distribution of some quantity of interest in each area, which contains more information than the area level aggregated data. For such data, we propose a new model-based approach for small area estimation. We assume that the unit level unobserved quantity of interest follows the linear mixed model. We develop an Monte Carlo EM algorithm to estimate the unknown parameters in the model and calculate the empirical best predictors of the area means by a simple Gibbs sampling algorithm. The numerical performance of our proposed method is examined by simulations. In addition, we apply the proposed method to the income dataset in Japan. In this application, we predict the Gini coefficient of each small area as well as mean income, which cannot be done by using Fay–Herriot model.

Keywords: Grouped data; Latent variables; Mixed effects model; Small area estimation.

---

*Graduate School of Social Sciences, Chiba University