Domingo Morales (Universidad Miguel Hernández de Elche)

Multivariate Fay–Herriot models for small area estimation

Topic 2 – Learning more from what we already know

Keywords: small area estimation, Multivariate linear mixed models, poverty mapping

Introduction

Multivariate Fay–Herriot models for estimating small area indicators are introduced. Among the available procedures for fitting linear mixed models, the residual maximum likelihood (REML) is employed. The empirical best predictor (EBLUP) of the vector of area means is derived. An approximation to the matrix of mean squared crossed prediction errors (MSE) is given and four MSE estimators are proposed. The first MSE estimator is a plug-in version of the MSE approximation. The remaining MSE estimators combine parametric bootstrap with the analytic terms of the MSE approximation.

Several simulation experiments are performed in order to assess the behavior of the multivariate EBLUP and for comparing the MSE estimators. The developed methodology and software are applied to data from the 2005 and 2006 Spanish living condition surveys. The target of the application is the estimation of poverty proportions and gaps at province level.

Methods / Problem statement

Surveys are designed for obtaining reliable estimates in the whole population or in some subpopulations called planned domains. However, it is quite common in practice to use survey data for estimating indicators of non-planned domains (small areas) with small samples sizes.

Results / Proposed solution

Small area estimation deals with inference problems for non-planned domains. In these cases, direct estimators might have large sampling errors. Direct estimators can be improved by assuming regression models that link all the sample data by introducing a relation between the variable of interest and a set of explanatory variables. This communication introduces model-based estimators based on multivariate Fay-Herriot models. Two applications are presented.

The target of the first application is the estimation of 2006 poverty proportions and gaps. The second application jointly estimates 2005 and 2006 poverty proportions.

Conclusions

This communication presents multivariate Fay–Herriot models for estimating small area parameters. Multivariate models incorporate the correlation of several target variables and borrow strength from auxiliary variables. The introduced models give some modeling flexibility, where they are applied to the estimation of poverty proportions and gaps in 2006 and to the estimation of poverty proportions in 2005 and 2006.

The communication gives the EBLUPs under the multivariate models and four MSE estimators. The presented simulations give some indications about the behavior of the EBLUPs and the MSE estimators.