Bivariate Small Area Estimation for Binary and Gaussian Variables Based on a Conditionally Specified Model

Hao Sun* Emily Berg*
Zhengyuan Zhu*
Department of Statistics, Iowa State University

Abstract

Many large-scale surveys collect both discrete and continuous variables. Small area estimates may be desired for means of continuous variables, proportions in each level of a categorical variable, or for domain means defined as the mean of the continuous variable for each level of the categorical variable. In this paper, we introduce a conditionally specified bivariate mixed-effect model for small area estimation, and provide a necessary and sufficient condition under which the conditional distributions render a valid joint distribution. The conditional specification allows better model interpretation. We use the valid joint distribution to calculate empirical Bayes predictors and use the parametric bootstrap to estimate the mean squared error. Simulation studies demonstrate the superior performance of the bivariate mixed-effect model relative to univariate model estimators. We apply the bivariate mixed-effect model to construct estimates for small watersheds using data from the Conservation Effects Assessment Project (CEAP), a survey developed to quantify the environmental impacts of conservation efforts. We construct predictors of mean sediment loss, the proportion of land where the soil loss tolerance is exceeded, and the average sediment loss on land where the soil loss tolerance is exceeded. In the data analysis, the bivariate mixed-effect model leads to more scientifically interpretable estimates of domain means than those based on two independent univariate models.

Keywords: Bivariate mixed-effect model; Bootstrap; EM algorithm; Small area estimation