Abstract

Transformations in the SAE Context
Comparing Generalized Linear Mixed Models and using an extension to Predictive Mean Matching Multiple Imputation of categorical variables Inference for the Potential Outcome framework Statistical misuse and consequences

Keywords

Multinomial Mixed Models in Small Area Estimation

The topic deals with small area estimation for the composition of categorical variables. A possible example is, for instance, the estimation of multidimensional poverty in small areas. From a methodological perspective, the thesis deals with small area estimation based on multinomial (mixed) models including the modeling of the mean squared error of the estimators. A simulation study could be conducted to investigate the properties of the estimators. Finally, the methods can be applied to poverty data from Mexico.

Multinomial Mixed Models in Small Area Estimation

Poverty Mapping using ELL-Methodology

The Quantile Regression (QR) is an extension of the multiple linear regression that specifies the change of conditional quantiles of a dependent variable, instead of changing the conditional mean. The additional information gained from this method might uncover new insights in datasets such as the one of the Demographic and Health Survey (DHS) of Senegal. QR could be used to estimate the conditional quantiles of wealth indicators such as the Wealth Index Factor Score for women and men, which uses the total value of an individual's assets to determine wealth, instead of commonly used monetary variables.

The Quantile Regression in Complex Design Samples

The robust Small Area Estimation Methods in the Context of Spatial Correlations

Robust Small Area Estimation Models in model-based settings

First, different methods (unit-level models under limited data set vs. area-level models [with and without aggregation]) will be compared in 4 model-based settings. If possible, it should then be investigated whether it might be useful to look for one of the methods is most convincing. A little application to real-world data is also possible.

Outliers and influential observations

Convergence in distribution is factually impossible to prove. The GO-diagnostic has become a widely accepted method to assess convergence in distribution. Investigate the Accuracy of the GO diagnostic using an extensive MC simulation study.

Logistic and multinomial models on the example of gender issues

Statistical misuse and consequences

Statistical literacy: Selection from typical problem clusters of statistical literacy, such as confusing, lack of measures of agranulosity and/or goodness of fit, correlation versus causality, relative versus absolute risk, various kinds of biases (e.g., self-selection, response bias, attrition, Hawthorne effect), social desirability bias, immortal time bias, etc.

Convergence in Distribution: Examining the Gelman-Rubin Diagnostic

Diagnostic for mass imputation: R package

Multivariate Imputation of categorical variables using an extension to Predictive Mean Matching

Comparing Generalised Linear Mixed Models and Transformations in the SAE Context