Author: Jay Breidt (Colorado State University)

Title: Model-Assisted Survey Estimation with Imperfectly Matched Auxiliary Data

Co-authors: Chien-Min Huang and Jean Opsomer

Abstract:

Model-assisted survey regression estimators combine auxiliary information available at a population level with complex survey data to estimate finite population parameters. Many prediction methods, including linear and mixed models, nonparametric regression, and machine learning techniques, can be incorporated into such model-assisted estimators. These methods assume that observations obtained for the sample can be matched without error to the auxiliary data. We investigate properties of estimators that rely on matching algorithms that do not in general yield perfect matches. We focus on difference estimators, which are exactly unbiased under perfect matching but not under imperfect matching. The methods are investigated analytically and via simulation, using a study of recreational angling in South Carolina to build a simulation population. In this study, the survey data come from a stratified, two-stage sample and the auxiliary data from logbooks filed by boat captains. Extensions to regression estimators under imperfect matching are discussed.