

Variance reduction using a non-informative sampling design

Thomas Zimmermann
Researcher
German Federal Statistical Office
<http://www.destatis.de>
Wiesbaden, Germany

Official Statistics commonly conducts sample surveys to produce estimates of aggregate statistics with a desired level of precision. For this purpose, design-based methods are used which are suitable for the estimation of finite population quantities such as totals or means. In most cases, however, model-based analyses are applied to the survey data as well. Examples include small area estimation techniques that allow for reliable estimates of finite population quantities in the presence of small sample sizes and socio-econometric models used in academia to test scientific hypotheses. This may cause problems as model-based methods frequently assume a non-informative sampling design and a violation of this assumption can lead to erroneous statistical inferences. We argue in this work that if the application of model-based methods can be anticipated before the sample is drawn, then this knowledge should be incorporated in the survey design. We propose a method called antithetic clustering that enables precise estimates for aggregate figures using design-based estimation methods and does automatically lead to non-informative sampling designs. We derive theoretical properties of this method under commonly employed models and conduct a simulation study to evaluate it numerically.