Small domain coverage estimation of the New Zealand 2018 Census: A case study in Bayesian analysis of complex survey data

Abstract

In many countries including New Zealand, a full census survey is the current gold standard to report official population statistics. As no census survey ever achieves 100% response rate, a smaller post-censal survey (PES) is usually conducted and analysed to assess census coverage and produce official population estimates by geographic area and demographic attributes such as age, sex, ethnicity, and place of birth. Considering the usually small size of the PES and the required level of disaggregation for population estimates, direct estimation is not feasible. Designed-based estimation including sampling weight adjustment is a commonly used method but is difficult to implement when survey non-response patterns and population benchmarks are not available. We overcome these limitations by developing and applying a fully model-based approach, using a Bayesian multilevel model. Although the theoretical bases for the Bayesian treatment of complex surveys have been described, published individual-level applications remain scarce. We provide such an application by using the 2018 census and PES surveys as a case study. We describe and implement a hierarchical model that accounts for the complex design of the PES. We then illustrate how mixed posterior predictive checking and cross-validation can assist with model building and model selection. Finally, we discuss potential methodological improvements and advantages and limitations of our approach in comparison to more traditional approaches.

Keywords