

Capture Recapture Methods for Hard-to-Count Population

Supervisor: [Prajamitra Bhuyan](#)

Research Group: [Statistics and Data Science](#)

Project description:

Capture-recapture methods are becoming an increasingly common approach for the study of under-reported events in the field of epidemiology, ecology, population studies, etc. Lincoln-Petersen estimate is popularly used under the assumption that the capture and recapture status of each individual is independent. However, this assumption often fails especially in the fields of public health and demography, which leads to bias in the Lincoln-Petersen estimate, and this is popularly known as correlation bias in the domain of census undercount estimation. Modeling of the capture-recapture data avoiding the causal independence assumption is a challenging problem in this domain.

International Working Group for Disease Monitoring and Forecasting presented a nice overview of the log-linear models for accounting dependency among the lists with heterogeneity in the capture probabilities. Chao and Tsay (1998) proposed a nonparametric population size estimator based on the sample coverage approach that incorporates the correlation bias for the three-sample model and measure of dependence between lists. Based on some simulation experiments, we observed that the coverage probability of the sample coverage-based estimate is very low. Moreover, this approach may provide an infeasible estimate in the presence of negative dependence among the lists. Fienberg et al. (1999) considered both classical and fully Bayesian hierarchical approaches to modeling individual heterogeneity and list interactions, which encompasses both the traditional loglinear approach and various elements from the full Rasch model. However, it is important to note that the parameters that are associated with the log-linear models are not well interpretable for practitioners.

The objective of this project to develop methods that successfully incorporates the possible causal dependency for the estimation of population size. The proposed model is applied to analyze real data sets from various scientific disciplines with a wide range of application areas, for example, the number of civilian casualties in a war or accident, human trafficking, drunk-drivers, and intravenous drug users.

References

Kiranmoy Chatterjee & Prajamitra Bhuyan (2020) On the estimation of population size from a post-stratified two-sample capture–recapture data under dependence, Journal of Statistical Computation and Simulation, 90:5, 819-838, DOI: 10.1080/00949655.2019.1707832

Further information:

[How to apply](#)

[Entry requirements](#)

[Fees and funding](#)