

Some new developments on likelihood approaches to estimation of participation probabilities for non-probability samples

Jean-François Beaumont

Statistics Canada

The estimation of finite population parameters using non-probability sample data is a research topic that has become popular in recent years. Often, methods that adjust non-probability sample estimates use auxiliary data collected in a probability sample and require the estimation of the probability of participation in the non-probability sample. In this presentation, we focus on parametric estimation of participation probabilities. First, we propose an improvement of the pseudo likelihood method of Chen, Li and Wu (2020), based on best linear unbiased estimation theory, that more efficiently leverages the available probability and non-probability sample data. We also develop a sample likelihood approach that properly accounts for the overlap between both samples when it can be identified in at least one of the samples. We use best linear unbiased prediction theory to handle the scenario where the overlap is unknown. Interestingly, our two proposed approaches coincide in the case of unknown overlap. Then, we show that many existing methods can be obtained as a special case of a general unbiased estimating function. Finally, we present the results of a simulation study that compares several methods.

Co-authors: Cynthia Bocci and Keven Bosa