

"Optimal" Calibration Weights Under Nonresponse in Survey Sampling

Per Gösta Andersson

Department of Statistics, Stockholm University

High nonresponse is a very common problem in sample surveys today. In statistical terms we are worried about increased bias and variance of estimators for population quantities such as totals and means. Different methods have been suggested in order to compensate for this phenomenon. We can roughly divide them into imputation and calibration and it is the latter approach we will focus on here.

In the nonresponse situation we can use linear calibration as in Särndal and Lundström (2005). The underlying distance function is not discussed in that paper, but it turns out that it corresponds to a GREG-type of distance function. The question arises if we can obtain better estimators by considering alternative distance functions. One candidate of such a function corresponds to the optimal regression estimator, see Andersson and Thorburn (2005). Under full response this estimator has asymptotically (in a design-based sense) minimum variance among regression types of estimators. This estimator is also a calibration estimator, thus fulfilling the calibration equation.

Under simple random sampling a GREG estimator and the optimal regression estimator are either identical or very close to each other. A simple sampling design for comparison where we achieve a substantial difference between the two types of estimators is Poisson sampling. Preliminary results from a simulation study generating nonresponse show that we can decrease both bias and variance using the alternative "optimal" weighting instead of the "standard" weighting.

References:

P.G. Andersson and D. Thorburn, An Optimal Calibration Distance Leading to the Optimal Regression Estimator, *Survey Methodology* 31 2005, 95-99.

CE Särndal, S. Lundström *Estimation in Surveys with Nonresponse*. Chichester, UK: Wiley, 2005.