Weight Smoothing via Design Modeling in Complex Surveys

F. Jay Breidt

NORC at the University of Chicago

In the classic design-based paradigm for complex surveys, study variables of interest are treated as nonrandom while each sample membership indicator is a binary random variable with known expectation, its first-order inclusion probability. The inverse of the inclusion probability is the design weight, interpretable as the number of population elements represented by the sampled element. We describe applications in which uncertainty in design categories (such as stratum or frame membership) leads to weight uncertainty and excess estimator variation, as well as making the representation interpretation less tenable. Each of these issues is addressed by smoothed design weights, related to those proposed by Beaumont (2008, Biometrika). The smoothed weights are obtained by modeling propensities for the uncertain categories using data external to the survey. The methodology is illustrated by modeling the propensity to be listed on a frame, leading to a smoothed dual frame estimator with application to surveys conducted by the National Agricultural Statistics Service.

Co-authors: Benjamin M. Reist