

JPSM/MPSM

Survey Methodology Seminar Series

Inferences for Respondent Driven Sampling

Z. TUBA SUZER-GURTEKIN

Presenter

Research Fellow, Institute for Social Research

Monday, April 20, 2015

12:00-1:00pm

Speaker will be at the University of Michigan, Video Room 368 - ISR South Basement
Room 1208 LeFrak Hall– Maryland

Abstract

RDS is a sampling and data collection method that has been widely employed in data collection for rare and hidden populations. However, the practices of RDS and underlying assumptions required for statistical inferences are yet to be understood clearly. In particular, the assumptions that govern the RDS inference are related to (1) Network structure, (2) Equilibrium condition, (3) Random recruitment, (4) Equal homophily, (5) Complete response, and (6) Accurate network size measures. Our ongoing research studies two particular aspects of the RDS inference: (1) the sampling process of RDS and how sampling error in RDS can be affected by other error sources, mainly nonresponse and measurement errors in relation to the RDS inference assumptions. We use total survey error framework to evaluate the implications of these errors theoretically and empirically, and (2) variance estimation. This presentation reports on the results from the second part, variance estimation. The RDS variance estimation is currently an open research area. Currently, two methods are implemented in the software: (1) Volz and Heckathorn closed-form (Volz and Heckathorn 2008), and (2) Bootstrapping (Salganik and Heckathorn 2007, Salganik 2006) variance estimators. First variance estimator requires to estimate the probability of a network member i of group A selecting a member j of the same group A , across all network members in the population. That is rather an unrealistic requirement given the limited information on the population network structure. Although the bootstrapping method does not require this information, both methods require static homophily across waves and across recruitment chains, equivalent to First Order Markov process. We propose an alternative bootstrapping variance estimator that relaxes the static homophily assumption. In our method, the resampling is applied to the seeds instead of relying on the sample derived group membership information as in the case of existing variance estimators. In a set of simulations that vary the selected number of seeds, we compared the 95% CI coverage of the existing and the proposed bootstrapping variance estimates. The simulations use the publicly available Sexual Acquisition and Transmission of HIV Cooperative Agreement Program (SATHCAP) data. Our results show that proposed method consistently yields higher coverage. In addition, small sample sizes, which are common in the RDS studies, present a problem for the existing bootstrapping method.