

## Respondent Driven Sampling as a Counting Process

Yakir Berchenko

Respondent driven sampling (RDS) is an approach to sample human populations that utilizes their social networks. Although RDS has become a widespread method for recruiting individuals within hidden populations, estimation using RDS is problematic due to biased sampling (e.g., over-sampling participants with many acquaintances). Most RDS studies attempt to adjust for this bias using inverse-degree weighting, assuming that the probability an individual is sampled is proportional to the number of their acquaintances (degree). However, this seemingly necessary assumption is unlikely to hold in practice. We propose a novel approach, that relaxes this assumption, by using a source of information that is usually ignored, the precise timing of recruitment. Our new approach, adapting methods developed for inference in epidemic processes, allows us also to test the assumption of recruitment proportional to degree, as well as to generate estimates of the total population size. We analyse these estimators and find them asymptotically consistent and normally distributed; applying them to five empirical RDS data-sets studied, we show that the probability an individual was sampled was not proportional to their degree, and in two datasets where the size and degree distribution can be characterised, we show that our new maximum likelihood estimator outperforms the standard inverse-degree weighting estimator.