

# **Statistics for cross-sectional surveys: Inferring total eventual time in current state using only elapsed time to date**

*Louis V. Cammarata*

*Student, Technology & Policy Program (IDSS)*

Using cross-sectional (“snapshot”) surveys, we wish to infer an individual’s or a system’s total time in a current temporary state based only on reported elapsed time to date. The survey question asked is: “How long have you been in this temporary state?” A recent paper by Larson (2016) developed a probability-based solution to this problem, utilizing properties of longevity bias. In this follow-up research, we investigate the statistical accuracy of the new method as a function of survey sample size, probability density function estimation technique, and properties of underlying distributions. We test several nonparametric estimation techniques, including kernel density estimation with data reflection, boundary-corrected kernel density estimation and non parametric maximum likelihood density estimation under shape constraints. The latter requires further processing as the estimator is inconsistent in 0. We report results utilizing Monte Carlo simulations with both discrete and continuous distributions for several types of sampling, including random and fixed times of surveying, as well as random and fixed times of entering the state. Bootstrap confidence interval and mean integrated squared error (MISE) results are provided. An example with NSF postdoc current duration data is included to demonstrate the steps.