

P8160 - Comparing Bootstrapping Methods Report

Amy Pitts, Hun Lee, Jimmy Kelliher, Tucker Morgan, Waveley Qiu

2/11/2022

Introduction

detail to be added

Simulation Planning

Aims

The primary goal of this simulation study is to assess the performance of two bootstrap methods (detailed below) in estimating the sampling variability of treatment effects obtained from a nearest-neighbor propensity-score matching (NNM). In this study, NNM will select a treated subject at random from simulated observational data. The untreated subject with the nearest propensity score is then selected to be paired with the treated subject, without replacement. Treatment effects can then be estimated by comparing outcomes (continuous or binary) between the treated and untreated subjects. The bootstrapping methods will be used to assess the estimates of variance of treatment effect.

Data Generation

The data for this simulation study will be generated from a parametric model. **Add details here**

Methods for Evaluation

Two bootstrap methods will be assessed in this simulation: the simple bootstrap and the complex bootstrap.

In the simple bootstrap, one draws repeated samples from the original sample with replacement to imitate the process of drawing repeated samples from a population. Here, repeated samples will be drawn from the matched pairs of observations. The standard deviation of the resulting samples will be used as an estimate of the standard error of the estimated treatment effect in the original sample of matched pairs.

The complex bootstrap considers two additional sources of variability compared to the simple bootstrap. In this approach, a sample is drawn with replacement from the original, unmatched observational data. The propensity-score model is estimated using this bootstrap sample, and NNM proceeds as before. The treatment effect is estimated from the newly matched sample.

The resulting standard error estimates will be the primary targets of this analysis.

Performance Measures

The standard error estimates from each bootstrap method will be assessed in two ways. First, standard error estimates will be compared to the standard deviation of treatment effect from the generated population to determine if the bootstrapping methods under- or over-estimated the parameter. Second, coverage rates of confidence intervals will be analyzed to assess how frequently the true average treatment effect is included in confidence intervals using the bootstrap-estimated standard errors.

Simulation Execution

discussion of coding will go here, perhaps even excerpts of code

Results

detail to be added

Discussion

conclusions to be added