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 variance of estimated treatment effects.

Data Generation

The data for this simulation study were generated from a parametric model. For each subject, three baseline covariates (L_1, L_2, L_3) were simulated from independent standard normal distributions. Two of these covariates $(L_1$ and $L_2)$ affected treatment selection, while two $(L_2$ and $L_3)$ affected the outcome. The probability of treatment for each subject was determined by the following model:

$$\text{logit}(\pi_i) = \alpha_0 + \alpha_1 L_{1i} + \alpha_2 L_{2i}$$

where α_0 was determined by:

$$
\begin{aligned}
E(T_i) &= E[E(T_i \mid \mathbf{X}_i, \boldsymbol{\alpha})] \\
&= E\left[\frac{\exp(\alpha_0 + \alpha_1 L_{1i} + \alpha_2 L_{2i} + \varepsilon_i)}{1 + \exp(\alpha_0 + \alpha_1 L_{1i} + \alpha_2 L_{2i} + \varepsilon_i)}\right] \\
&= \frac{\exp(\alpha_0 + \alpha_1 E(L_{1i}) + \alpha_2 E(L_{2i}) + E(\varepsilon_i))}{1 + \exp(\alpha_0 + \alpha_1 E(L_{1i}) + \alpha_2 E(L_{2i}) + E(\varepsilon_i))} \\
&= \frac{\exp(\alpha_0)}{1 + \exp(\alpha_0)} \\
\implies \pi_0 &= \frac{\exp(\alpha_0)}{1 + \exp(\alpha_0)} \\
\implies \alpha_0 &= \frac{\log \pi_0}{\log(1 - \pi_0)}
\end{aligned}
$$

for a given π_0.

For each type of outcome (continuous and binary) and each bootstrapping method (simple and complex), 100 sub-populations of 5,000 subjects will be generated using the following parametric model:

$$Y_i = \beta_o + \beta_1 A_i + \beta_2 L_{2i} + \beta_3 L_{3i} + \epsilon_0$$

where Y_i indicates the binary or continuous outcome for each subject, A_i indicates the treatment status of each subject (0 or 1), and L_{2i} and L_{3i} indicate covariate values for each subject.
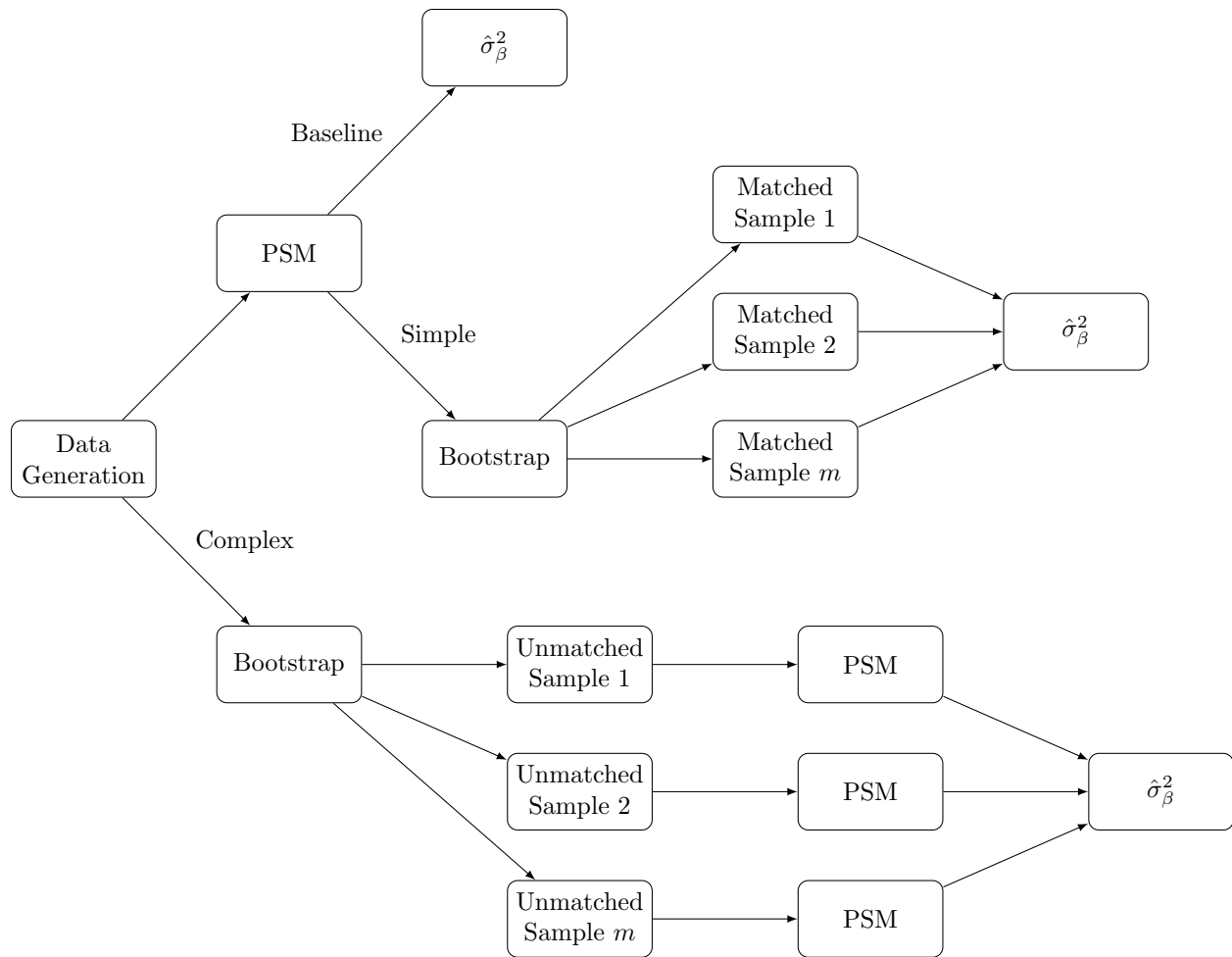
More details. . .

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 an original sample with replacement to imitate the process of drawing repeated samples from a population. Here, 1,000 repeated samples (m_{boot}) of 500 pairs (n_{boot}) will be drawn from the matched pairs of observations for each of the 100 sub-populations (m_{sample}). *The standard deviation of the resulting estimated treatment effects will be used as an estimate of the standard error of the true treatment effect.*

```
# Do I have the above italicized section and parameter specifications correct?
```

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. This process is repeated 1,000 times (m_{boot}) for each of the 100 sub-populations (m_{sample}).

The resulting standard error estimates, σ_β^2, will be the primary targets of this analysis.

$\hat{\sigma}_\beta^2$

Baseline

PSM

Matched Sample 1

Matched Sample 2

$\hat{\sigma}_\beta^2$

Simple

Data Generation

Bootstrap

Matched Sample m

Complex

Bootstrap

Unmatched Sample 1

PSM

Unmatched Sample 2

PSM

$\hat{\sigma}_\beta^2$

Unmatched Sample m

PSM

Performance Measures

The standard error estimates from each bootstrap method will be assessed in two ways. First, 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. Second, standard error estimates from each bootstrap method will be compared to the sample standard deviation of treatment effects of the generated sub-populations to determine how bootstrapping aligns with a simpler approach.

Simulation Execution

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

Results

detail to be added

Discussion

conclusions to be added