

Does your alternative matching method have more runs with higher proportions of balanced covariates?

Your Answer: In the simulation of 100 runs using the alternative matching method, the proportion of covariates that met the balance threshold ($|SMD| \leq 0.1$) ranged from 0.000 to 1.000, with a median of 0.600 and a mean of 0.636. Out of 1,000 simulations, 45% of the runs produced a proportion of balanced covariates above the median. This variability suggests that while many runs achieve decent balance, there is still a substantial degree of fluctuation in the covariate balance across different model specifications. Such variability is expected due to random covariate selection; however, the fact that only 45% of runs exceed the median indicates that the method does not consistently produce high balance.

Use a visualization to examine the change in the distribution of the percent improvement in balance in propensity score matching vs. the distribution of the percent improvement in balance in your new method. Which did better? Analyze the results in 1-2 sentences.

Your Answer: The distribution of the mean percent improvement in balance for the alternative matching method shows a median improvement of 67.51% and a mean of 71.55%, with values ranging from 44.05% to 100%. When compared to the standard propensity score matching results, the alternative method seems to yield a stronger and more consistent improvement in balance on average. However, the variability observed indicates that some model specifications still lead to lower improvements in balance, highlighting the potential sensitivity of the matching results to covariate selection.

Discussion Questions

Why might it be a good idea to do matching even if we have a randomized or as-if-random design?

Your Answer: Even in randomized or as-if-random designs, matching can further refine the balance between treatment and control groups by reducing any residual imbalances that occur by chance. This additional step can lead to more precise estimates of the treatment effect and improve the credibility of the causal inference by ensuring comparability between groups.

The standard way of estimating the propensity score is using a logistic regression to estimate probability of treatment. Given what we know about the curse of dimensionality, do you think there might be advantages to using other machine learning algorithms (decision trees, bagging/boosting forests, ensembles, etc.) to estimate propensity scores instead?

Your Answer: Yes, alternative machine learning algorithms, such as decision trees, random forests, or boosting methods, can capture complex, non-linear relationships among covariates better than logistic regression. In high dimensional settings, these methods may reduce bias and improve covariate balance more effectively, leading to more robust treatment effect estimates.