Causal Inference Algorithms

Project4-Group3

Xiangning Han, Charles Shin, Sneha Swati, Shuyuan Wang, Minzhi Zhang

Agenda

1. Overview
2. Propensity Score Estimation
3. Algorithm Overview
4. Evaluation and Comparison
5. Q & A

1

Overview

Goal

- Implementing, evaluating and comparing three algorithms for Causal Inference (i.e. estimating true ATE).
- The algorithms implemented are Stratification, Regression Adjustment and Weighted Regression, together with L2 penalized logistic regression as the propensity score estimation method.
- Performed the ATE estimation on two generated datasets, low dimensional and high dimensional dataset.

2

Propensity Score Estimation

Propensity Score Estimation

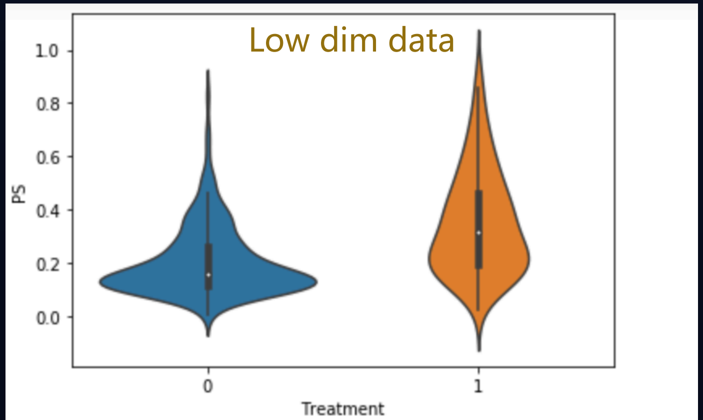
L2 penalized logistic regression as the propensity score estimation method.

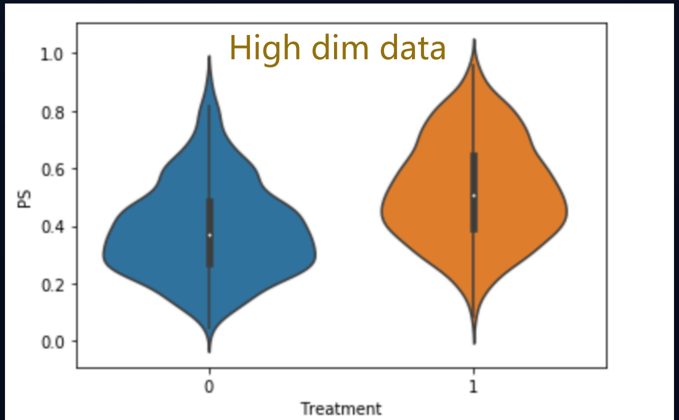
The propensity score is defined by: e(x)=Pr(T=1|X=x)

L2 penalized logistic regression: introducing regularization term to decrease the model variance in the loss function Q. Modifying the loss function with a penalty term, which is 'L2 norm'

$$Q = -\frac{1}{n}\sum_{i=1}^{n}[y_i(\beta_0 + \beta_0 x_{i1.} + \cdots + \beta_{ip}) + \log(1 + \exp(\beta_0 + \beta_0 x_{i1.} + \cdots + \beta_{ip}))] + \lambda \sum_{j=1}^{p} \beta_j^2$$

For both low dimensional data and high dimensional data, with λ (parameter for the penalty term) chosen as 2, the violin plots show that the distribution of propensity scores for both groups are similar, indicating an acceptable overlapping.





3

Algorithm Overview

Stratification

Data	Run Time	ATE Est
Low dim data	0.030781	2.879272
High dim data	0.013375	-2.783537

Stratification on the propensity score involves stratifying subjects into mutually exclusive groups based on their estimated propensity score.

Divided subjects into five equal-size groups by the quintiles of the estimated propensity score.

The average of all the difference of average outcome between treated and untreated group for each stratum is the estimated ATE.

Regression Adjustment

Data	Run Time	ATE Est
Low dim data	0.033268	2.502966
High dim data	0.130357	-2.999067

Remove the effects of confounding when estimating the effects of treatment on outcomes via covariate adjustment using the estimated propensity score.

The outcome variable Y is regressed on the indicator variable A(T) and the estimated propensity score (PS).

Used a linear regression model. The estimated coefficient on the indicator variable A(T) is the estimated ATE.

Weighted Regression

Data	Run Time	ATE Est
Low dim data	0.136385	2.523955
High dim data	0.685741	-2.964053

Weighted least square estimation of the regression function:

$$Y_i = \alpha_0 + \tau * T_i + \alpha_1' * T_i + \alpha_2' * (Z_i - \bar{Z}) * T_i + \varepsilon_i$$

The outcome variable Y was regressed on the selected covariates Z.

To select Z, estimated K linear regressions of the type :
$Y_i = \beta_0 + \beta_{k1} * T_i + \beta_{k2} * X_{ik} + \varepsilon_i$, where K is the total number of covariates.

Calculate the t-statistic for the test of the null hypothesis that the slope coefficient $\beta_{k2} = 0$ in each of these regressions, selected for Z all the covariates with a t-statistics larger in absolute value than $t_{reg} = 2$.

The estimated coefficient of the indicator variable T $\hat{\tau}$ is the estimated ATE.

4

Evaluation and Comparison

Evaluation and Comparison

Evaluation method:

Absolute error and relative error to evaluate the ATE estimates

Absolute error(ABSE) is given by:

$$ABSE = |ATE_{EST} - ATE_{True}|$$

Relative error is given by:

$$RE = |ATE_{EST} - ATE_{True}| / |ATE_{EST}|$$

Evaluation and Comparison

Low dim data				
Algorithm	Run Time	ATE Est	Abs Error	Relative Error
Stratification	0.030781	2.879272	0.38	15.17%
Regression Adjustment	0.033268	2.502966	0.00	0.12%
Weighted Regression	0.136385	2.523955	0.02	0.96%

High dim data				
Algorithm	Run Time	ATE Est	Abs Error	Relative Error
Stratification	0.013375	-2.783537	0.22	7.22%
Regression Adjustment	0.130357	-2.999067	0.00	0.03%
Weighted Regression	0.685741	-2.964053	0.04	1.20%

Evaluation and Comparison

Conclusion:

- Regression Adjustment is the best method for both low dimensional data and high dimensional data.
- Stratification has a fastest run time among the three Algorithms.

5

Q & A

Reference

- Austin, Peter C. 2011. “An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies.” Multivariate Behavioral Research 46 (3): 399–424.
- Hirano, Keisuke, and Guido W Imbens. 2001. “Estimation of Causal Effects Using Propensity Score Weighting: An Application to Data on Right Heart Catheterization.” Health Services and Outcomes Research Methodology 2 (3-4): 259–78.
- Lunceford, Jared K, and Marie Davidian. 2004. “StratiÞcation and Weighting via the Propensity Score in Estimation of Causal Treatment Effects a Comparative Study.” Statistics in Medicine 23 (19): 2937–60.
- Westreich, Daniel, Justin Lessler, and Michele Jonsson Funk. 2010. “Propensity Score Estimation: Neural Networks, Support Vector Machines, Decision Trees (Cart), and Meta-ClassiÞers as Alternatives to Logistic Regression.” Journal of Clinical Epidemiology 63 (8): 826–33.

Thank you!