**Asymmetry of Doubly Robust Estimators**

**1. Introduction**

The double robust estimators (DR) are a class of causal effect estimators which corrects bias in the causal effect estimated using propensity score model or outcome model. The double robust estimators remain asymptotically unbiased if one of the two models is correctly specified and ensure consistency even if one model is wrong. Theoretically, double robust estimator postulates that propensity score model and outcome model play a symmetric role in ensuring consistency of double robust estimators. If either of the models is correct, it corrects the bias in causal estimate to ensure its consistency given large sample size. This project aims to investigate the conventionally assumed symmetric role of propensity score model and outcome model in consistency of double robust estimator (DR). We investigated this assumed symmetric role of PS model and outcome model considering different sample sizes and dimensions. In this study, we considered various ways to construct double robust estimators and compared performance of double robust estimators and non-double robust estimators based on simulated datasets of different sizes and dimensions. Comparing performance of these DR estimators using correctly specified and incorrectly specified models helps us to look at consistency of DR estimators with respect to sample size and dimensions in comparison to non-double robust estimators. The intellectual ideas to specify model misspecification for DR estimators and to generate simulation dataset have been borrowed from significant work done by Kang and Schafer [1]. The assumptions and methodology considered to generate the simulated datasets and estimate causal effect using DR estimators is mentioned below.

**2. Modeling and Data generation**

For i=1,2,…n

W=

is 4\*4 identity matrix.

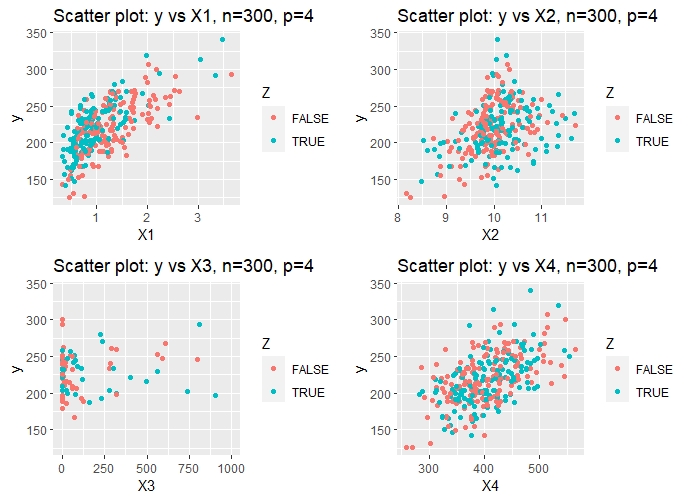
Generate true propensity

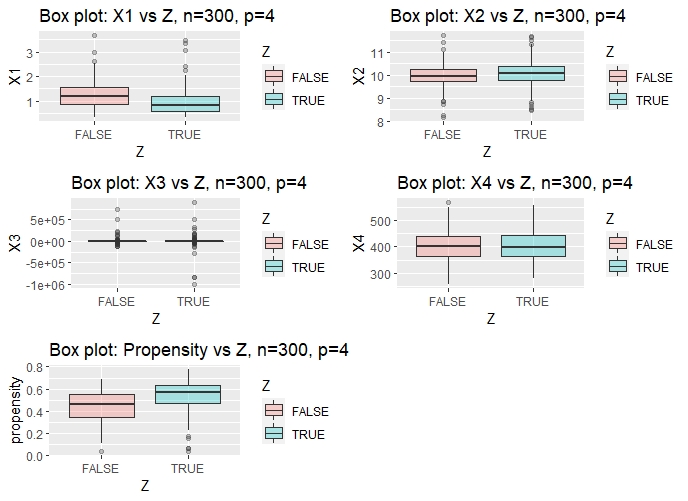
Generate true causal y1 and y0

Scatter plot for covariate and y

Box plot of ps for treatment/control

Other causal assumptions – show graphs





**3. Estimating ATE on simulated data**

After simulating the datasets which satisfy fundamental assumptions for causal inference, we estimated causal effect using Double Robust estimators based on combination of correctly specified and incorrectly specified propensity model and outcome model for various sample sizes and dimensions. Finally we compared performance of these DR estimators with correct/incorrect propensity model and potential outcome model to draw performance comparison inference.

Below mentioned are the correct/incorrect propensity model and potential outcome model used to estimate ATE (average treatment effect)

Correct PS model:

Incorrect PS model:

The ATE was estimated using normalized IPW (inverse propensity weighed) estimate based on correct/incorrect PS model

Correct PO model:

Incorrect PO model:

Double Robust Estimators:

We initially considered un-normalized version of double robust (DR) estimator for inference. Since the un-normalized version of DR estimator had high variance especially when propensity and outcome model both are incorrectly specified, we resorted to normalized version of DR estimator as suggested by Kang and Schafer.

Un-normalized DR Estimator:

Normalized DR Estimator:

The above mentioned correctly/incorrectly specified models led us to estimate eight causal effects which are presented below:

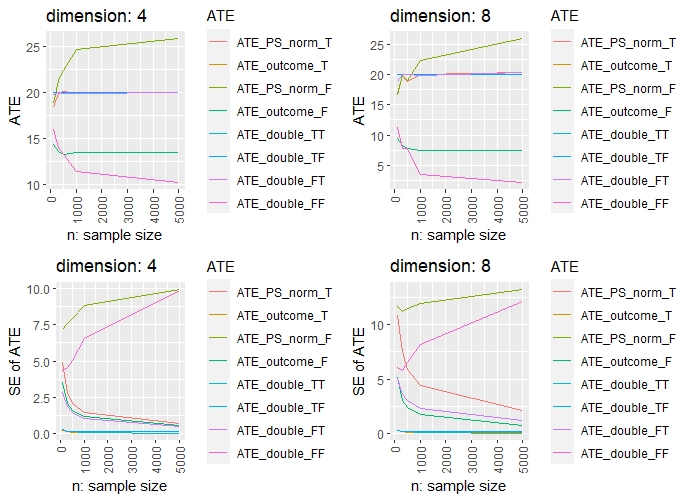
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ATE Estimator | Propensity Model (Correctly Specified) | Potential Outcome Model (Correctly Specified) | Reference Name | |
| Normalized IPW | Yes |  | | ATE\_PS\_norm\_T |
| Normalized IPW | No |  | | ATE\_PS\_norm\_F |
| OLS estimate |  | Yes | | ATE\_outcome\_T |
| OLS estimate |  | No | | ATE\_outcome\_F |
| Double Robust | Yes | Yes | | ATE\_double\_TT |
| Double Robust | Yes | No | | ATE\_double\_FT |
| Double Robust | No | Yes | | ATE\_double\_TF |
| Double Robust | No | No | | ATE\_double\_FF |

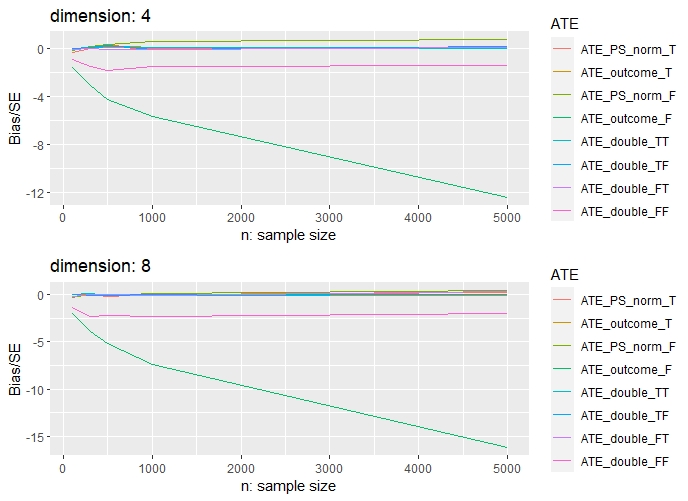
The simulation exercise carried out to estimate these causal effects included 100 simulations for each causal estimate. Each simulation generated datasets of different sizes (100, 300, 500, 1000, 5000) and dimensions (4 and 8) which were used to estimate the ATE. Each simulation included standard error estimation based on 200 bootsrap samples.

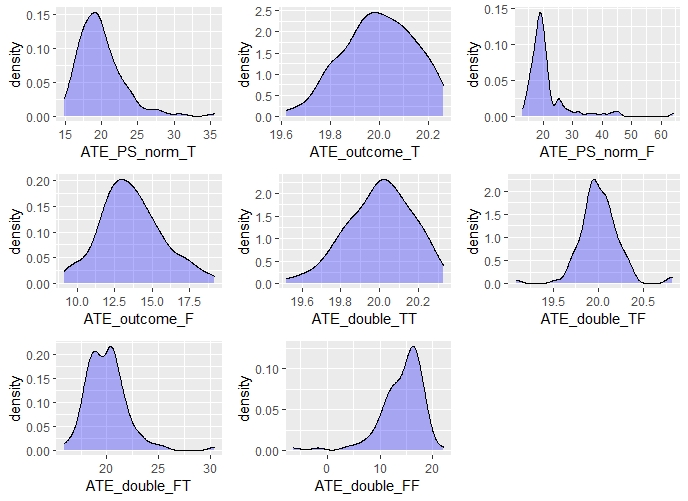
**4. Performance of DR estimators on simulated data**

First, we repeated the simulation with sample sizes of 100, 300, 500, 1000, and 5000. This was primarily to check the rates of convergence. The first graph plots ATE's for each estimator that are averaged over 100 simulations, and the second graph depicts their standard errors, also averaged across simulations. The bias can be read off from the first graph. As expected, every estimator which has at least one model correctly specified exhibits fast convergence to the true mean, 20. On the other hand, the bias of is more than and. So the use of doubly robust estimator comes with a caveat that one should be sufficiently confident with at least one of the models.

Now we move on to the second graph. As expected, the standard errors of, and are very small from the beginning. They are followed by , and whose standard errors decreases close to 1 with the sample size of 1000. This implies that their rate of convergence is moderately fast. and, on the other hand, do not exhibit decreasing trend. On the contrary, their standard errors increase to 10 even with the sample size of 5000. Looking through their ATE estimates, we could find that this was due to a very few extreme valued ATE estimates. Thus, it is advisable to check for extremely small propensity scores or compare mean absolute error (MAE) with mean squared error (MSE) as the former is more robust to the outliers. This shows that some mis-specification of the propensity score model could produce an extreme ATE estimate when it is principally dependent on propensity scores. Interestingly, does not exhibit this extreme behavior even though its propensity score model is misspecified in the same way.







**5. Conclusion**

Bibliography

1. [Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data](https://arxiv.org/abs/0804.2958)