Annals of Studies in Science and Humanities
Volume 1 Number 2 (2015) : 53–59

http://journal.carsu.edu.ph/
Online ISSN 02408-3631

Efficiency of Doubly Robust Estimator of Causal Effects with Trimmed and Winsorized Propensity Score Weights

Angela Fel A. Padecio[1,*] *and Carolina B. Baguio*[2]

[1]Department of Mathematics, Caraga State University, Butuan City, Philippines
[2]Department of Statistics, Mindanao State University - Iligan Institute of Technology

Received: November 3, 2014 Accepted: June 7, 2016

ABSTRACT

Estimation of treatment effect with causal interpretation where treatment is not randomized may be biased if confounding is not taken into appropriate account. Adjustment for confounding is often carried out through regression adjustment (ANCOVA) or propensity score (PS) method specifically inverse weighting. When used individually to estimate a causal effect, both ANCOVA and PS method are unbiased only if the model is correctly specified. The doubly protected or doubly robust (DR) estimator combines these two approaches such that only one of the models need be correctly specified to obtain an unbiased effect estimator. This paper focuses on the scenario with correct ANCOVA model but misspecified PS model. PS weighting is sensitive to model misspecification and outlying weights that can excessively influence results. To deal with this problem, robust methods are used to lessen the influence of extreme weights through trimming and winsorization. A Monte Carlo simulation framework is used based on real-world data modeling the use of statin to lower the cholesterol. Empirical results show that trimming the PS weights gives an improvement on the DR estimator efficiency in terms of its root mean square error (RMSE) only for small sample sizes. On the other hand, winsorizing the PS weights considerably improves the efficiency of DR estimator both for small and large samples.

Keywords: Doubly Robust Estimation, Regression Adjustment (ANCOVA), Propensity Score, Trimming, Winsorization

* Corresponding Author
Email:afpadecio@gmail.com

1 Introduction

Estimation of the effect of a treatment with causal interpretation from studies where treatment is not randomized may be biased if selection effect such as confounding is not taken into appropriate account. Adjustment for confounding is often carried out through regression adjustment (ANCOVA) or propensity score (PS) method specifically inverse weighting. Historically, regression adjustment (ANCOVA) has been used more regularly than any other methods for estimating the effects of treatments when using observational data. Alternative methods for control of confounding in observational studies based on the propensity score (PS) were proposed by Rosenbaum and Rubin (1983). When used individually to estimate a causal effect, both ANCOVA and PS methods are unbiased only if the statistical model is correctly specified. However, Kreif et al. (2012) found that most studies use regression or PS weighting but do not cautiously evaluate model specification.

Robins et al. (1994) introduced the notion of doubly protected or doubly robust (DR) estimator that combines these two approaches such that only one of the two models need be correctly specified to obtain an unbiased effect estimator. One possible scenario of misspecification is when the ANCOVA model is correctly specified but the PS model is misspecified. This results to extremely large PS weights that may improperly influence the outcomes and this would result in estimates with high variance. One probable solution is to apply the methods of robust statistics to lessen the influence of extreme weights through trimming and winsorization. Baguio (1999) showed that the optimum amount of trimming is 5% to 10% while Yusof et al. (2011) found that the most ideal total amount of trimming is 15%. Bollinger and Chandra (2005) found that winsorizing is almost always better than trimming and winsorizing at 1% is preferred to doing nothing. They added that 5% is better than 1% winsorization.

This study compares the efficiency of DR estimator when PS weights are trimmed and untrimmed and DR estimator when PS weights are winsorized and unwinsorized.

This provides an analyst the opportunity for obtaining unbiased inference when adjusting for selection effects such as confounding by agreeing for diverse forms of model misspecification such as regression model that is correctly specified with PS model misspecified. This study, furthermore, will distinguish the potential improvement of the DR estimator after trimming or winsorizing the extreme PS weights due to misspecified PS model. Moreover, this is a potentially valuable tool for obtaining more robust effect estimates in observational studies of the effects of drugs, devices and other interventions.

2 R Algorithm on DR Estimator with Untrimmed or Trimmed and Unwinsorized and Winsorized PS Weights

A Monte Carlo simulation framework introduced by Setoguchi et al. (2008) is used which is based on real-world claims data modeling the use of a drug called statin that can lower the cholesterol. The ranges of variables ($\alpha_i's$ and $\beta_i's$) used in simulations are

consistent with the expected ranges in the actual pharmacoepidemiology practice.

Datasets are generated with $n = 100$, 300 and 500 observations with binary treatment T, continuous outcome Y and 10 covariates $X_i, i = 1, ..., 10$. Four of $X_i(X_i, i = 1, ..., 4)$ are independently associated with both T and Y, which are then referred to as the confounders, three of $X_i(X_i, i = 5, 6, 7)$ are associated with the treatment only, referred to as the exposure predictors, and three of $X_i(X_i, i = 8, 9, 10)$, called the outcome predictors, are associated with the outcome only. Bootstrap samples of the same size are then generated for each of these sample sizes. The true PS model which has a moderate non-additivity and non-linearity property defines the association between the exposure and covariates which is based on realistic scenario. Datasets are generated 500, 1000, and 1500 times.

The following is a modified algorithm based from Lee et al. (2011):

Step 1: Generate the 10 $X_i's$ and ϵ as independent standard normal random variables with zero mean and unit variance of sample size 100.

Step 2: Model the binary treatment, T, using logistic regression as a function of X_i, for $i = 1, \ldots, 7$.

The true PS model is

$$\begin{aligned} P[T = 1|X_i] =& 1 + \exp[-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 \\ &+ \beta_7 X_7 + \beta_2 X_2^2 + \beta_4 X_4^2 + \beta_7 X_7^2 + 0.5\beta_1 X_1 X_3 + 0.7\beta_2 X_2 X_4 \\ &+ 0.5\beta_3 X_3 X_5 + 0.7\beta_4 X_4 X_6 + 0.5\beta_5 X_5 X_7 + 0.5\beta_1 X_1 X_6 \\ &+ 0.7\beta_2 X_2 X_3 + 0.5\beta_3 X_3 X_4 + 0.5\beta_4 X_4 X_5 + 0.5\beta_5 X_5 X_6)]^{-1} \end{aligned}$$

where $\beta_0 = 0, \beta_1 = 0.8, \beta_2 = -0.25, \beta_3 = 0.6, \beta_4 = -0.4, \beta_5 = -0.8, \beta_6 = -0.5, \beta_7 = 0.7$.

Generate a random number between 0 and 1 from a uniform distribution. T is set to be 1 if the randomly generated number is less than the estimated true PS, and 0 otherwise.

Step 3: Model the outcome Y as a function of $X_i, i = 1, \ldots, 4, 8, 9, 10$ and T, that is,

$$\begin{aligned} E(Y|T, X_1, X_2, X_3, X_4, X_8, X_9, X_10) =& \alpha_0 + \theta T + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 \\ &+ \alpha_4 X_4 + \alpha_5 X_8 + \alpha_6 X_9 + \alpha_7 X_1 0 + \epsilon \end{aligned}$$

where $\alpha_0 = -3.85, \alpha_1 = 0.3, \alpha_2 = -0.36, \alpha_3 = -0.73, \alpha_4 = -0.2, \alpha_5 = 0.71, \alpha_6 = -0.19, \alpha_7 = 0.26, \theta = -0.4$.

Step 4: Generate a bootstrap sample of the same size.

Step 5: Fit a misspecified PS model using logistic regression for treatment T conditional on $X_i, i = 1, ..., 7$ with missing interactions and nonlinearities, that is, predicted values from this regression give the estimated PS, $\hat{e}_i$.

Step 6: Calculate the PS weights, that is, $\frac{1}{\hat{e}_i}$ if the subject receives the treatment and $\frac{1}{1-\hat{e}_i}$ if the subject receives the control, where $\hat{e}_i$ is the estimated PS.

Step 7: Fit the regression model specified in Step 3 for the treatment group only ($T = 1$), and obtain the predicted values for the whole sample. This gives the value for $m_1(\mathbf{X}, \hat{\alpha}_1)$. Also, fit the outcome regression model on the said variables for the control group only ($T = 0$), and obtain the predicted values for the whole sample. This gives the value for $m_0(\mathbf{X}, \hat{\alpha}_0)$.

Step 8: Substitute the PS weights, $m_0(\mathbf{X}, \hat{\alpha}_0)$ and $m_1(\mathbf{X}, \hat{\alpha}_1)$ into the expression for the DR estimator.

Step 9: Repeat these steps 500, 1000 and 1500 times and compute for the mean, variance, bias and RMSE of the DR estimator for each bootstrap on different replication.

Step 10: The high estimated PS weights from Step 6 are trimmed downwards, with percentile cutpoints ranging from the 99th to the 50th percentiles, at 1% intervals.

Step 11: Repeat Step 7 to Step 9. Compare the relative efficiency (RE) of the DR estimator with untrimmed PS weights and DR estimator with trimmed PS weights.

Step 12: Using the same misspecified PS model in Step 5, the high estimated PS weights in Step 6 are now winsorized, with percentile cutpoints also ranging from the 99th to the 50th percentiles, at 1% intervals.

Step 13: Repeat Step 7 to Step 9. Compare the relative efficiency (RE) of the DR estimator with unwinsorized PS weights and DR estimator with winsorized PS weights.

Do these steps also for sample sizes $n = 300$ and $n = 500$ with the same number of bootstrap replications $B = 500, 1000$ and 1500 for each sample size.

3 Summary of Findings

Investigations have been made for increasing sample size $n = 100$, 300 and 500 for different bootstrap replications $B = 500$, 1000 and 1500 to determine the behavior of estimators particularly DR estimator with PS weights which are neither trimmed nor winsorized, DR estimator with trimmed PS weights and DR estimator with winsorized PS weights. Results from Table 1 show that for smaller sample sizes $n = 100$ and 300, the RMSE of the DR estimator with trimmed PS weights is more efficient than the DR estimator with untrimmed PS weights upon trimming small proportions of the PS weights. However, for larger sample size $n = 500$, DR estimator with untrimmed PS weights is more efficient. The results further show that these hold true in different bootstrap replications aforementioned.
Moreover, DR estimator with winsorized PS weights in Table 2 found to be more efficient than that of the estimator with unwinsorized weights valid for all winsorizing proportions for sample sizes $n = 100$ and 300. This holds true in all mentioned bootstrap replications. For larger sample size $n = 500$, DR estimator with winsorized PS weights dominates over the estimator with unwinsorized weights for greater than 8%, 1% and 2% PS weight winsorization for bootstrap replications $B = 500$,1000 and 1500, respectively.

4 Conclusion

Empirical results show that trimming and winsorizing the PS weights substantially give an improvement on the efficiency of the doubly robust estimator of causal effects in view of some cases and considerations. DR estimator with trimmed PS weights is observed to dominate in terms of efficiency over the DR estimator with untrimmed weights for small trimming proportions valid for smaller sample sizes and holds true for different bootstrap replications. However, this result does not apply for larger sample sizes implying that the DR estimator with untrimmed weights is preferable. This indicates that the asymptotic efficiency of the DR estimator with trimmed PS weights does not hold true. On the other hand, DR estimator with winsorized PS weights performs better than the DR estimator with unwinsorized weights for small trimming proportions which remains true for all sample sizes in different bootstrap replications. This suggests that winsorizing the weights considerably improves the efficiency of DR estimator.

5 Recommendations

The following recommendations are made regarding future work on the doubly robust (DR) estimator, particularly on trimming or winsorizing the propensity score (PS) weights.

1. The study uses only one choice of parameter values specified in the work of Setoguchi et al (2008) typical in pharmacoepidemiologic studies which is based on real-world claims data modeling the use of a drug to lower the cholesterol called statin. Other range of parameter values is suggested.

2. The study makes use of a propensity score (PS) model with moderate non-additivity and non-linearity property. It is advised to work with a model having other than the above-mentioned property such as additivity and linearity, mild non-linearity, moderate non-linearity, mild non-additivity, mild non-additivity and non-linearity, and moderate non-additivity properties.

3. Determine the performance of DR estimator when PS model is correctly specified but misspecified regression model.

4. Determine the causal effects of binary treatment on a binary outcome, continuous treatment on a binary outcome and continuous treatment on a continuous outcome.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

Baguio, C. (1999). *Adaptive Robust Estimators of a Location Parameter. Doctoral Dissertation.* (Unpublished doctoral dissertation). MSU-Iligan Institute of Technology, Iligan City, Philippines.

Bollinger, C. R., and Chandra, A. (2005). Iatrogenic specification error: a cautionary tale of cleaning data. *Journal of Labor Economics, 23*(2), 235-257.

Kreif, N., Grieve, R., and Sadique, M. Z. (2013). Statistical methods for cost-effectiveness analyses that use observational data: A critical appraisal tool and review of current practice. *Health economics, 22*(4), 486-500.

Lee, B. K., Lessler, J., and Stuart, E. A. (2011). Weight trimming and propensity score weighting. *PloS one, 6*(3), e18174.

Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American statistical Association, 89*(427), 846-866.

Rosenbaum, P. R., and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 41-55.

Setoguchi, S., Schneeweiss, S., Brookhart, M. A., Glynn, R. J., and Cook, E. F. (2008). Evaluating uses of data mining techniques in propensity score estimation: a simulation study. *Pharmacoepidemiology and drug safety, 17*(6), 546-555.

Yusof, Z., Abdullah, S. and Yahaya, S. S. S. (2011). Type I error rates of F_t statistic with different trimming strategies for two groups case. *Modern Applied Science, 5*(4), 236-242.

Table 1: RMSE of Doubly Robust Estimator with Trimmed and Untrimmed Propensity Score Weights

Trimming %	$B = 500$			$B = 1000$			$B = 1500$		
	$n = 100$	$n = 300$	$n = 500$	$n = 100$	$n = 300$	$n = 500$	$n = 100$	$n = 300$	$n = 500$
0	0.7643	0.3515	0.1972	0.7429	0.3404	0.1984	0.7466	0.3388	0.1984
1	0.7472	0.2792	0.2020	0.7302	0.2721	0.2015	0.7337	0.2692	0.2009
2	0.7381	0.2887	0.2135	0.7227	0.2818	0.2115	0.7289	0.2787	0.2107
3	0.7349	0.2959	0.2164	0.7186	0.2885	0.2161	0.7261	0.2855	0.2149
4	0.7329	0.3012	0.2179	0.7168	0.2948	0.2179	0.7259	0.2928	0.2159
5	0.7357	0.3100	0.2185	0.7205	0.3030	0.2179	0.7289	0.3022	0.2164
6	0.7393	0.3169	0.2200	0.7230	0.3104	0.2199	0.7328	0.3093	0.2180
7	0.7468	0.3225	0.2223	0.7288	0.3142	0.2227	0.7394	0.3131	0.2217
8	0.7551	0.3276	0.2254	0.7356	0.3203	0.2265	0.7466	0.3198	0.2258
9	0.7642	0.3313	0.2319	0.7447	0.3233	0.2313	0.7548	0.3236	0.2306
10	0.7708	0.3405	0.2362	0.7540	0.3312	0.2365	0.7640	0.3311	0.2362
11	0.7764	0.3477	0.2414	0.7605	0.3377	0.2411	0.7696	0.3373	0.2417
12	0.7821	0.3545	0.2472	0.7676	0.3450	0.2459	0.7770	0.3429	0.2469
13	0.7874	0.3626	0.2502	0.7741	0.3500	0.2485	0.7842	0.3477	0.2495
14	0.7940	0.3689	0.2540	0.7824	0.3558	0.2517	0.7928	0.3526	0.2526
15	0.8015	0.3738	0.2551	0.7883	0.3607	0.2538	0.7992	0.3582	0.2551
16	0.8128	0.3777	0.2582	0.7954	0.3649	0.2555	0.8050	0.3631	0.2566
17	0.8213	0.3832	0.2607	0.8009	0.3691	0.2576	0.8100	0.3672	0.2580
18	0.8271	0.3856	0.2610	0.8059	0.3728	0.2584	0.8159	0.3701	0.2589
19	0.8315	0.3892	0.2631	0.8107	0.3772	0.2608	0.8209	0.3746	0.2612
20	0.8362	0.3915	0.2636	0.8151	0.3805	0.2612	0.8243	0.3773	0.2616

Table 2: RMSE of Doubly Robust Estimator with Winsorized and Unwinsorized Propensity Score Weights

Winsorization %	$B = 500$			$B = 1000$			$B = 1500$		
	$n = 100$	$n = 300$	$n = 500$	$n = 100$	$n = 300$	$n = 500$	$n = 100$	$n = 300$	$n = 500$
0	0.7643	0.3516	0.1972	0.7429	0.3404	0.1984	0.7466	0.3388	0.1984
1	0.7097	0.2744	0.2010	0.6860	0.2740	0.1997	0.6918	0.2698	0.2001
2	0.7032	0.2729	0.1999	0.6794	0.2715	0.1983	0.6825	0.2675	0.1987
3	0.7010	0.2724	0.1993	0.6777	0.2707	0.1976	0.6820	0.2663	0.1978
4	0.6988	0.2720	0.1988	0.6768	0.2697	0.1971	0.6810	0.2652	0.1971
5	0.6975	0.2721	0.1984	0.6765	0.2690	0.1962	0.6809	0.2645	0.1971
6	0.6960	0.2717	0.1981	0.6757	0.2685	0.1958	0.6804	0.2640	0.1971
7	0.6957	0.2718	0.1978	0.6759	0.2682	0.1954	0.6806	0.2637	0.1952
8	0.6953	0.2717	0.1974	0.6755	0.2678	0.1950	0.6804	0.2635	0.1948
9	0.6945	0.2717	0.1970	0.6754	0.2677	0.1946	0.6805	0.2634	0.1944
10	0.6941	0.2716	0.1968	0.6752	0.2675	0.1943	0.6804	0.2632	0.1941
11	0.6933	0.2716	0.1966	0.6747	0.2673	0.1941	0.6802	0.2631	0.1939
12	0.6930	0.2714	0.1964	0.6747	0.2671	0.1939	0.6801	0.2630	0.1936
13	0.6929	0.2715	0.1962	0.6746	0.2670	0.1937	0.6800	0.2630	0.1934
14	0.6922	0.2715	0.1961	0.6742	0.2670	0.1936	0.6796	0.2630	0.1932
15	0.6919	0.2716	0.1959	0.6740	0.2670	0.1934	0.6794	0.2630	0.1930
16	0.6918	0.2716	0.1958	0.6739	0.2669	0.1932	0.6793	0.2630	0.1928
17	0.6915	0.2717	0.1957	0.6738	0.2668	0.1932	0.6792	0.2629	0.1926
18	0.6913	0.2716	0.1956	0.6737	0.2667	0.1930	0.6791	0.2628	0.1925
19	0.6910	0.2717	0.1955	0.6736	0.2667	0.1929	0.6791	0.2628	0.1923
20	0.6908	0.2717	0.1954	0.6734	0.2667	0.1928	0.6790	0.2628	0.1922