Comparison of Conventional Propensity Score Approaches with Bayesian Propensity Score in continuous data: A Simulation study

302376

OakPil Han[1], Hyori Yun[1]
[1]Hanmi Pharm. Co., Ltd, Seoul, South Korea

ABSTRACT

The propensity score has become an acceptable and powerful tool to eliminate imbalance in the distribution of confounding variables between treatment groups and quite often applied for observational studies or real world data analysis these days. We conducted Monte Carlo simulation for continuous data to investigate the performance of matching on the propensity score, stratification on the propensity score and covariate adjustment using the propensity score for continuous outcomes by 4 different set of variables for the propensity score model.

Similar to the previous work by Austin[1,2] for binary outcomes, we generated 9 independent variables with different association strength with the treatment and the outcome. The simulation studies explored 4 conventional propensity score models by different choices of variables: 1) all variables associated with treatment allocation 2) all variables associated with Outcome 3) true confounder 4) all variables. Each propensity score model was assessed by sample mean difference and mean squared error (MSE). In addition, we applied a two-step bayesian propensity score analysis (BPSA) and compared with the conventional approach[3].

We found that matching on the PS resulted in the smallest MSE followed by adjustment for PS and PS based stratification in the ability of treatment effect estimation. However, there was no significant difference between conventional PS and Bayesian PS based methods across different PS models.

BACKGROUND

There are many literatures showing adjusted indirect comparison using propensity score-based approach for the studies which can not apply randomized controlled trials. The propensity score is well known as the probability of treatment assignment conditioning on the observed baseline covariates. The propensity score can help facilitate the dimension reduction to adjust for covariate imbalance. Thus, it would be crucial to determine to which variables to include in the propensity score model. Austin et al (2007) investigated very comprehensively different scenarios about covariate inclusion for binary outcome[1]. Further to that, we tried to investigate similar scenarios in continuous outcome by a series of Monte Carlo simulations.

Within the perspective of propensity score, the score is dependent on the observed covariates included in the model. In the simulations, we assumed different fixed associations for the covariates but in reality, it is not easy to know the true association. In consideration of propensity score is a probability and uncertainty in model parameters of the propensity score model, we adopted bayesian perspective computing bayesian propensity score with the same data generated for the simulations and compared the results with those by conventional propensity score analysis (PSA)[3].

METHODS

Conventional Propensity Score

1. Generate independent variables by different association[1]
   Draw each of nine independent variables, xi ~ N(0,1) in different association with the treatment and outcome under sample size n=100 per arm as below:

		Association with treatment		
		Strong	Moderate	None
Association with outcome	Strong	X_1	X_2	X_3
	Moderate	X_4	X_5	X_6
	None	X_7	X_8	X_9

2. Treatment assignment based on the independent variables:
   E(logit(P_i , treatment)) = $\beta_{0,treat} + \sum_{i=1}^{9} \beta_i X_i$, where i =1,2,...,9
   $treatment_i \sim$ Bernoulli(P_i)
   Following the same scenarios of Austin *et al.*[1], the regression coefficients in the logit model above were set as follows:
   $\beta_1, \beta_3, \beta_5$: log(5)
   $\beta_2, \beta_4, \beta_6$: log(2)
3. Generate an outcome conditional on the treatment and assumed variables with different association:
   $E(Y_i) = \alpha_0 + \alpha_{treat} T_i + 5 X_1 + 5 X_2 + 5 X_3 + 2 X_4 + 2 X_5 + 2 X_6$
   , where $\alpha_0 = 30$, $\alpha_{treat} = 0$
4. Replicate the above steps 100 times

Bayesian Propensity Score

1) Used the same simulation data generated for the conventional PS models above
2) Assume prior distributions for $\beta_{0,treat}$ and $\beta_i's (i = 1,2,\dots,9)$ for the PS model as follows:

$\beta_{0,treat} \sim N(0, B^{-1})$
$\beta_1 \sim N(\log 5, B^{-1})$
$\beta_2 \sim N(\log 2, B^{-1})$
$\beta_3 \sim N(0, B^{-1})$
$\beta_4 \sim N(\log 5, B^{-1})$
$\beta_5 \sim N(\log 2, B^{-1})$
$\beta_6 \sim N(0, B^{-1})$
$\beta_7 \sim N(\log 5, B^{-1})$
$\beta_8 \sim N(\log 2, B^{-1})$
$\beta_9 \sim N(0, B^{-1})$
B = 0 (Non-informative) ,1 ,10 ,100

3) Estimate mean posterior coefficients of the PS model by MCMClogit function in R with a thining interval of 10 after 50,000 burn-in in 100,000 iterations.
4) Estimate Bayesian PS using the coefficients from 3)

Different Propensity Score models by variables included

PS model	Independent variables included in the PS model
PS1	$X_1, X_2, X_4, X_5, X_7, X_8$
PS2	$X_1, X_2, X_3, X_4, X_5, X_6$
PS3	X_1, X_2, X_4, X_5
PS4	X_1 - X_9 (All variables)

For the different PS model types, we computed the mean differences between treatment groups along with the mean squared error (MSE) in order to simply quantify the ability of PS models as below.

- MSE for treatment effect = $((\bar{X}_{treatment} - \bar{X}_{control}) - 0)^2$
  , where
  $\bar{X}_{treatment}$, $\bar{X}_{control}$ = mean of the variable in treated and control group
  true effect size was assumed as 0 in the simulation

RESULTS

We tried to quantify the ability of PS models by mean difference estimation between treatment and control group and MSE.
The first comparison for the four PS models by different approaches is presented in the Figure 1 below. The PS based matching appears to be closest to true effect size, 0, followed by covariate adjustment for PS and stratification method.
Also, we investigated Bayesian PS(BPS) for the same four PS models, PS1 – PS4 using the same simulation data to compare with conventional PS approach. The PS based matching approach was compared with BPS based matching by different prior variance as Figure 2. The BPSA with non-informative prior (BPSA_0) appeared to be very close to the conventional PSA line for the PS1 and PS2 models but not for PS3 or PS4.

Figure 1. Conventional PS models by different method

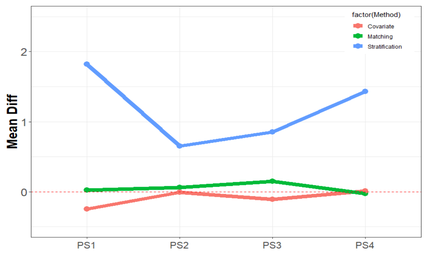
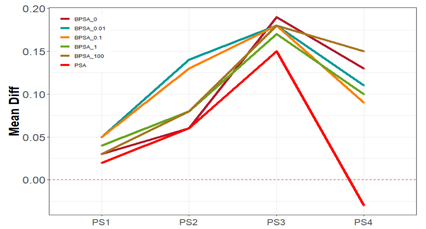


Table 1. Mean diff and MSE by different PS models

		PS based Matching		PS as Covariate		PS based Stratification	
		Mean diff	MSE	Mean diff	MSE	Mean diff	MSE
PSA	PS1	0.02	2.65	-0.25	2.98	1.82	5.91
	PS2	0.06	0.87	-0.01	1.64	0.65	4.39
	PS3	0.15	1.59	-0.11	1.57	0.85	4.43
	PS4	-0.03	1.67	0.01	3.09	1.43	5.88
BPSA (Non-informative)	PS1	0.03	2.67	-0.08	3	0.98	4.77
	PS2	0.06	0.92	0.05	1.65	0.81	3.26
	PS3	0.19	1.59	-0.09	1.57	0.78	3.2
	PS4	0.13	1.67	0.12	3.12	1	4.89
BPSA (B^{-1}=0.01)	PS1	0.05	2.79	-0.32	2.86	0.17	4.69
	PS2	0.14	0.95	0.03	1.55	0.67	3.1
	PS3	0.18	1.71	0.05	1.53	0.69	3.09
	PS4	0.11	1.51	-0.33	2.89	0.21	4.7
BPSA (B^{-1}=0.1)	PS1	0.05	2.73	-0.33	2.95	0.39	4.71
	PS2	0.13	0.94	-0.07	1.62	0.67	3.2
	PS3	0.18	1.7	-0.1	1.56	0.7	3.14
	PS4	0.09	1.73	-0.27	3.06	0.47	4.8
BPSA (B^{-1}=1)	PS1	0.04	2.64	-0.16	2.99	0.93	4.75
	PS2	0.08	0.9	0.03	1.65	0.81	3.26
	PS3	0.17	1.62	-0.1	1.57	0.77	3.2
	PS4	0.1	1.83	0.07	3.13	1.1	4.84
BPSA (B^{-1}=100)	PS1	0.03	2.67	-0.09	3	0.9	4.76
	PS2	0.08	0.93	0.05	1.65	0.83	3.27
	PS3	0.18	1.59	-0.09	1.57	0.79	3.19
	PS4	0.15	1.69	0.2	3.14	1.01	4.88

Figure 2. PSA vs. BPSA by different prior variance for Matching method



CONCLUSION

Based on the simulated sample mean difference between groups and the mean squared error by different PS based approach,

1) Overall, matching on the PS method resulted in the smallest MSE followed by adjustment for PS and PS based stratification in ability of true effect size estimation
2) Within PS based matching approach, PS2 model with the variables associated with outcome resulted in the smallest MSE among the 4 different PS models
3) There was no significant difference in mean difference or MSE between conventional PS and Bayesian PS based methods across 4 PS models (PS1-PS4) and 3 methods (matching, covariate, stratification) in continuous data
4) In Bayesian PS based analyses, the estimation for the treatment effect difference between groups did not appear to be impacted much by prior precision (non-informative, B=0.01,0.1, 1,100)

CONSIDERATION

It would be good if we had a real case study to apply the investigated methods.

There could be further quantification to investigate the ability of PS based methods in balance for the independent variables after matching or stratification method though we did not present herein.

Also, it would be good if further research could investigate the present uncertainty in the results by
- taking various number of sample size per group; and
- increasing the number of repetition in the simulation ; and
- applying various distribution of independent variables; and
- applying hierarchical bayesian approach for the posterior distribution of model coefficients

REFERENCES

1. Austin PC, Grootendorst P, Anderson GM. A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a Monte Carlo study. Stat Med. 2007 Feb 20;26(4):734-53. doi: 10.1002/sim.2580. PMID: 16708349.
2. Austin PC. Some methods of propensity-score matching had superior performance to others: results of an empirical investigation and Monte Carlo simulations. Biom J. 2009 Feb;51(1):171-84. doi: 10.1002/bimj.200810488. PMID: 19197955.
3. Kaplan, D., Chen, J. A Two-Step Bayesian Approach for Propensity Score Analysis: Simulations and Case Study. Psychometrika 77, 581–609 (2012). https://doi.org/10.1007/s11336-012-9262-8