Propensity Score Matching in Observational studies: A case study

Usha K Thamattoor

The goal of propensity scoring is to mimic randomized controlled trials by balancing observed covariates between subjects in control and treatment study groups

Aim

- A propensity score is the conditional probability that a subject receives "treatment" given the subject's observed covariates. The goal of propensity scoring is to mimic what happens in randomized controlled trials (RCT's) by balancing observed covariates between subjects in control and treatment study groups (Faries, Leon, Haro, Obenchain, 2010). A caliper width defines the range within which the propensity scores (or logit of the propensity scores) must fall to be considered a valid match
- **Objective**: The poster will walk through a case study to explain propensity score matching (PSM) using SAS® to match samples employing nearest neighbour pair-matching without replacement with a caliper width of 0.2 (Austin, PC 2011).

Methods

- Data was simulated to implement Propensity score matching.
- Propensity scores, the probability to be two treatment groups control and treated subjects (1, 0) was calculated for each participant conditional on the following variables collected at baseline:
  - Age, type of Diabetes, duration of Diabetes Mellitus, Body mass index(BMI), tobacco use, alcohol use
- The nearest neighbor pair-matching without replacement with a caliper width of 0.2 of the standard deviation of the logit of the propensity score was used as shown in Fig1. Selected baseline variables was compared between the two treatment groups (1, 0) before and after propensity score matching.

Methods

- **SAS Code**
  - Logistic regression was used to create propensity scores

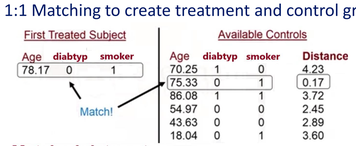
```
Proc logistic data=datain;
class  var1 var2 /param=reference ;
model treatment (event="1") = var1 var2 var3 var4/ link=logit; output
out=DataPSscore prob=pscore; run;
```

  - Nearest neighbor pair-matching without replacement with a caliper width of 0.2

```
retain BestDistance 99;
rc=iter.first();   if (rc=0) then BestDistance= 99;
do while (rc= 0);     if (pscoreT - 0.2) < pscoreC < (pscoreT + 0.2) then do;
ScoreDistance= abs(pscoreT - pscoreC);     if ScoreDistance < BestDistance
then do;
BestDistance= ScoreDistance;   IdSelectedControl= csubjid;
MatchedToTreatID= tsubjid;   end;  end;
rc = iter.next(); if (rc ~= 0) and BestDistance~= 99 then do;
if (rc ~= 0) and BestDistance~= 99 then do;   output;   rc1= h.remove(key:
IdSelectedControl);   end;  end;
```

The estimated propensity score is saved in the variable pscore and it is saved, along with all the covariates, in the dataset called DataPscore

Fig1: 1:1 Matching to create treatment and control groups



Key results

Table 1: Before propensity score matching(N=499)

	Before propensity score matching(N =499)		
	Treated subjects 1	**Control 0**	P-value
Age (years)			
N	305	194	0.9273
Mean (SD)	32.84 (5.66)	32.97 (6.26)	
Median	31	32	
2.5; 97.5 Percentiles	21.00 ; 45.00	21.00 ; 45.00	
Min ; Max	19.00 ; 45.00	18.00 ; 45.00	
Duration of DM (years)			
N	305	194	<0.001
Mean (SD)	12.36 (7.89)	16.59 (8.53)	
Median	10	16	
2.5; 97.5 Percentiles	2.00 ; 27.00	2.00 ; 32.00	
Min ; Max	1.00 ; 31.00	1.00 ; 36.00	
Body Mass Index (kg/m2)			
N	305	194	0.0217
Mean (SD)	23.52 (5.22)	23.61 (3.77)	
Median	22.4	23.1	
2.5; 97.5 Percentiles	19.00 ; 39.70	18.70 ; 31.90	
Min ; Max	18.00 ; 52.50	17.60 ; 46.40	
Type of diabetes, N (%)			
N	305	194	0.2506
TYPE 1	237 (77.7)	142 (73.2)	
TYPE 2	68 (22.3)	52 (26.8)	
Tobacco, N (%)			
N	303	194	0.5368
CURRENT SMOKER	38 (12.5)	19 (9.8)	
PREVIOUS SMOKER	24 (7.9)	13 (6.7)	
NEVER SMOKED	241 (79.5)	162 (83.5)	
Alcohol, N (%)			
N	303	193	0.5229
Yes	11 (3.6)	5 (2.6)	
No	292 (96.4)	188 (97.4)	

Table 2: After propensity score matching (N=328)

	After propensity score matching(N =328)		
	Treated subjects 1	**Control 0**	P-value
Age (years)			
N	164	164	0.28
Mean (SD)	32.41 (4.63)	31.84 (5.84)	
Median	31	31	
2.5; 97.5 Percentiles	25.00 ; 45.00	21.00 ; 45.00	
Min ; Max	19.00 ; 45.00	18.00 ; 45.00	
Duration of DM (years)			
N	164	164	0.0002
Mean (SD)	17.94 (6.27)	15.04 (8.07)	
Median	19	16	
2.5; 97.5 Percentiles	5.00 ; 29.00	3.00 ; 31.00	
Min ; Max	2.00 ; 31.00	2.00 ; 36.00	
Body Mass Index (kg/m2)			
N	164	164	0.6191
Mean (SD)	24.22 (6.01)	23.47 (3.90)	
Median	22.7	23.1	
2.5; 97.5 Percentiles	19.40 ; 41.20	18.70 ; 31.60	
Min ; Max	18.00 ; 52.50	17.60 ; 46.40	
Type of diabetes, N (%)			
N	164	164	1
TYPE 1	116 (70.7)	116 (70.7)	
TYPE 2	48 (29.3)	48 (29.3)	
Tobacco, N (%)			
N	164	164	0.8447
CURRENT SMOKER	17 (10.4)	19 (11.6)	
PREVIOUS SMOKER	11 (6.7)	13 (7.9)	
NEVER SMOKED	136 (82.9)	132 (80.5)	
Alcohol, N (%)			
N	164	164	1
Yes	5 (3.0)	5 (3.0)	
No	159 (97.0)	159 (97.0)	

Differences between treatment groups are tested using chi-square test for categorical variables and Wilcoxon rank test for continuous variables, N: Number of patients, SD: Standard deviation. Percentages are based on total number of patients in the analysis set with information.

- 499 subjects were simulated. Baseline characteristics were compared before and after PSM. All the characteristics age, type of Diabetes, BMI, tobacco use, alcohol use were similar between treatment groups (1, 0) after propensity score matching except duration of Diabetes Mellitus which is associated with treatment (1,0) (Ref. Table 1 and Table 2)

Conclusion

- Most of the bias influencing the treatment was removed using the propensity score methodology

Novo Nordisk Service Centre India Private Ltd

References:

(**1**) Roy, Jason. A Review of: Analysis of Observational Health Care Data Using SAS, by DE Faries, AC Leon, JM Haro, and RL Obenchain (Eds.).Cary, NC: SAS Institute, Inc., 2010. ISBN 978-1-60764-227-5, xiv+ 436 pp., $64.95. (2010): 173-174.

(**2**) Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharm Stat. 2011 Mar; 10(2): 150–161.