

Section 10: Inverse propensity score weighting (IPSW)

Yotam Shem-Tov

Fall 2014

Inverse Propensity Score Weighting (IPSW)

- Until now we discussed matching on the P-score, a different approach is to re-weight the observations by the P-score
- By weighting the observations according to the true P-score, we can make the treated and control observation have exactly the same distribution of covariate,

$$F_{X|T=1}(x) = F_{X|T=0}(x)$$

where $x = (x_1, x_2, \dots, x_p)$, and p is the number of covariates

- This means that weighting on the P-score has clear testable implications
- After weighting all the moments of the joint distribution of X should be the same between the treatment and the control groups

Inverse Propensity Score Weighting (IPSW)

- Is the identifying assumptions in IPSW different than matching (with or without replacement)?
- No, the assumptions are the same:
 - 1 Conditional Independence Assumption (CIA)
 - 2 Overlap condition
 - 3 SUTVA
- IPSW is a smooth estimator, this allows us to use bootstrap to calculate the SE of the treatment effect estimator, accounting for the variance due to the P-score estimation procedure
- [Abadie and Imbens \(2008\)](#) show the failure of the bootstrap in matching estimators.
- Sub-sampling bootstrap (m out of n) works well for estimating the SE of matching estimators.
- IPSW is frequently used in quantile regressions, when adjusting for X . See Bitler, Gelbach, and Hoynes (2005) for an example

Inverse Propensity Score Weighting (IPSW)

- Denote by $p_i(x)$ the true propensity score, and by $\hat{p}_i(x)$ the estimated propensity score
- The IPSW weights are,

$$w_i = \begin{cases} \frac{1}{p_i(x)}, & T_i = 1 \\ \frac{1}{1-p_i(x)}, & T_i = 0 \end{cases} \Leftrightarrow w_i = \frac{T_i}{p_i(x)} + \frac{1 - T_i}{1 - p_i(x)}$$

- Consider the estimator

$$\hat{\tau}_{ipsw} = \frac{\sum_{i=1}^N Y_i T_i}{\sum_{i=1}^N \frac{T_i}{p_i(X)}} - \frac{\sum_{i=1}^N Y_i (1 - T_i)}{\sum_{i=1}^N \frac{1 - T_i}{1 - p_i(X)}}$$

- See [Hirano and Imbens \(2001\)](#) for a discussion of this estimator and properties
- This is the weighted difference in means after adjusting the weights to sum to 1
- Is this estimator unbiased with respect to *ATT*? How about with respect to *ATE*?

Inverse Propensity Score Weighting (IPSW)

- Assume we know the *true* propensity score, and that CIA hold, $(Y_1, Y_0) \perp T|X$
- Is it correct that?

$$(Y_1, Y_0) \perp T|p(X)$$

Yes , why? By Rosenbaum and Rubin (1983), the propensity score is also a balancing score, and by *Theorem 3* in the paper, it follows that the potential outcomes are independent of the treatment conditional on a balancing score, $b(x)$, and assuming $0 < P(T = 1|b(x)) < 1$

- Is the following claim correct

$$\mathbb{E} \left(\frac{Y \cdot T}{p(X)} \right) = \mathbb{E}(Y_1) \text{ ?}$$

The law of iterated expectations (LIE),

$$\mathbb{E}_F(F) = \mathbb{E}_G[\mathbb{E}_{F|G}(F|G)|G]$$

Note that, for any random variable A and indicator variable T :

$$\begin{aligned}\mathbb{E}[A \cdot T] &= \mathbb{E}[A \cdot 1] \cdot P(T = 1) + \mathbb{E}[A \cdot 0] \cdot (1 - P(T = 1)) \\ &= \mathbb{E}[A] \cdot P(T = 1)\end{aligned}$$

Next we prove the claim:

$$\begin{aligned}\mathbb{E} \left(\frac{Y \cdot T}{p(X)} \right) &= \mathbb{E} \left[\mathbb{E} \left(\frac{Y \cdot T}{p(X)} | X \right) \right] \\&= \mathbb{E} \left[\mathbb{E} \left(\frac{Y \cdot T}{p(X)} | X, T = 1 \right) P(T = 1 | X) \right] \\&= \mathbb{E} \left[\mathbb{E} \left(\frac{Y_1}{p(X)} | X, T = 1 \right) p(X) \right] = \mathbb{E} \left[\mathbb{E}(Y_1 | X, T = 1) \frac{1}{p(X)} \cdot p(X) \right] \\&= \mathbb{E} [\mathbb{E}(Y_1 | X, T = 1)] = \mathbb{E} [\mathbb{E}(Y_1 | X)] = \mathbb{E}(Y_1)\end{aligned}$$

Using the same line of proof we can show that,

$$\mathbb{E} \left(\frac{Y \cdot (1 - T)}{1 - p(X)} \right) = \mathbb{E}(Y_0)$$

- Therefore $\hat{\tau}_{ipsw}$ is an unbiased estimator for the ATE
- In order for $\hat{\tau}_{ipsw}$ to be an unbiased estimator for the ATE we needed to assume strong ignorability, i.e
 $0 < P(T = 1|b(x)) < 1$
- In order to estimate ATT , we can relax the overlap condition to, $0 < P(T = 1|b(x)) \leq 1$
- In order to estimate ATC , we can relax the overlap condition to, $0 \leq P(T = 1|b(x)) < 1$
- See Imbens (2000) for a generalization of the IPSW estimator for the case of multi-valued treatments. In this scenario the treatment variable is an integer which can take $\{0, 1, \dots, K\}$ different values

Example: Lalonde (1986), Dehejia and Wahba (1999)

- Lalonde (1986) analysed the effect of a randomized control trial of a job training program, the NSW data
- Using the experimental data Lalonde can estimate well both $\mathbb{E}(Y_1)$ and $\mathbb{E}(Y_0)$.
- Hence, we can estimate ATE ,

$$\mathbb{E}(Y_1) - \mathbb{E}(Y_0) = \mathbb{E}(Y_1 - Y_0) = ATE$$

- *Lalonde compare the estimate of $\mathbb{E}(Y_0)$ using NSW, to estimators based on observational data, using CPS and PSID samples*
- Lalonde used the extraordinary data to evaluate how good are estimators for $\mathbb{E}(Y_0)$, based on observational data

Example: Lalonde (1986), Dehejia and Wahba (1999)

- Dehejia and Wahba (1999) used this data to argue in support of propensity score matching estimation
- They argued the propensity score matching (using observational data) can recover the experimental estimation, and started a long discussion and exchange of articles for many years
- We will not take a stand in the argument, we will use this data to compare between matching, IPSW, and regression adjustment
- The data we will use is from Dehejia and Wahba (1999), which restricted the original NSW data to individuals with information on prior earnings, before the job training program (earnings in 1974, 1975)
- Link to the NSW data is [here](#), link to the CPS-1 data is [here](#)

Exmple: Lalonde (1986), Dehejia and Wahba (1999)

NSW controls vs NSW treatment

	Ave. Treat	Ave. control	T-test	Wilcoxon	KS
age	25.816	25.054	0.266	0.215	0.748
ed	10.346	10.088	0.150	0.056	0.063
black	0.843	0.827	0.647	0.649	1.000
hisp	0.059	0.108	0.064	0.077	0.963
married	0.189	0.154	0.334	0.327	0.999
nodeg	0.708	0.835	0.002	0.001	0.063
re74	2095.574	2107.027	0.982	0.361	0.970
re75	1532.055	1266.909	0.385	0.061	0.164

Exmple: Lalonde (1986), Dehejia and Wahba (1999)

CPS-1 data vs NSW controls

	Ave. NSW	Ave. CPS	T-test	Wilcoxon	KS
age	25.054	33.225	0.000	0.000	0.000
ed	10.088	12.028	0.000	0.000	0.000
black	0.827	0.074	0.000	0.000	0.000
hisp	0.108	0.072	0.067	0.028	0.901
married	0.154	0.712	0.000	0.000	0.000
nodeg	0.835	0.296	0.000	0.000	0.000
re74	2107.027	14016.800	0.000	0.000	0.000
re75	1266.909	13650.804	0.000	0.000	0.000

Exmple: Lalonde (1986), Dehejia and Wahba (1999)

CPS-1 controls vs NSW treated

	Ave. Treat	Ave. control	T-test	Wilcoxon	KS
age	25.816	33.225	0.000	0.000	0.000
ed	10.346	12.028	0.000	0.000	0.000
black	0.843	0.074	0.000	0.000	0.000
hisp	0.059	0.072	0.475	0.510	1.000
married	0.189	0.712	0.000	0.000	0.000
nodeg	0.708	0.296	0.000	0.000	0.000
re74	2095.574	14016.800	0.000	0.000	0.000
re75	1532.055	13650.804	0.000	0.000	0.000

Exmple: Lalonde (1986), Dehejia and Wahba (1999)

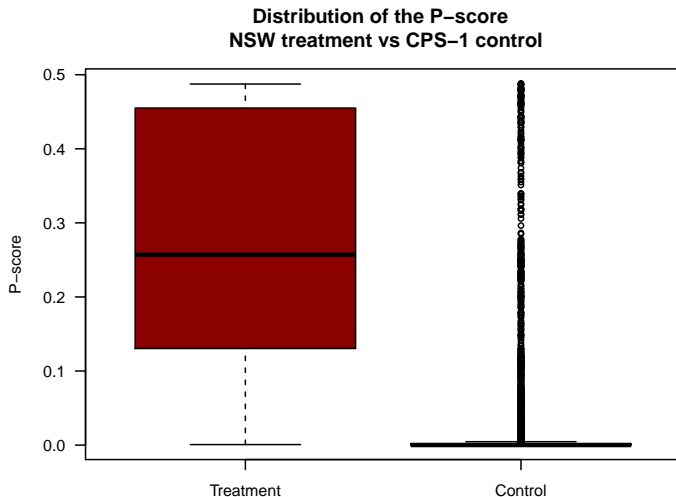
- Compare NSW treatment with CPS-1 control using propensity score (hence PS)
- We estimate PS using a simple additive logit model,

```
x1 = d1[,c("age","ed","black","hisp","married","nodeg",  
"re74","re75")]
```

```
# P-score estimation
```

```
ps.model1 < glm(treat~(.),  
data=data.frame(treat=d1$treat,x1),  
family=binomial(link=logit))
```

Exmple: Lalonde (1986), Dehejia and Wahba (1999)



	Regression NSW	Regression CPS-1	IPSW	Matching
treat	1675.86** (639.34)	793.59 (548.25)	-1263.46*** (136.54)	985.86 (782.81)
age	141.73 (274.42)	-233.68*** (41.18)	-421.47*** (41.83)	234.64 (324.96)
ed	385.02 (230.11)	166.85*** (28.66)	-15.29 (31.91)	368.07 (230.59)
R ²	0.06	0.48	0.45	0.22
Adj. R ²	0.03	0.48	0.45	0.22
Num. obs.	445	16177	16177	9748

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

- The matching is with replacement using only the P-score,

`match=Match(Tr=treat1,X=ps,replace=TRUE,ties=FALSE)`

Should we include ties? **Yes**

- The IPSW yields an extremely biased estimator, can we say why? Is it our fault?

IPSW has testable implications and diagnostics we can do in order to check if the identifying hold and whether something "went bad" in the estimation

1 Check the distribution of the weights

```
> summary(w[treat1==0])
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 1.000   1.000   1.000   1.012   1.002   1.955

> summary(w[treat1==1])
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 2.052   2.198   3.894  45.320   7.667 1425.000
```

If $\hat{e}(x) \approx 1$ or $\hat{e}(x) \approx 0$, what will happen to the weights?

This is the case when the overlap condition does not hold

2 Look on a balance table using the weighted means and weighted t-tests

The balance table when weighting using the P-score

	Ave. Treat	Ave. control	T-test
age	26.921	33.135	0.000
ed	11.402	12.008	0.000
black	0.268	0.082	0.000
hisp	0.098	0.072	0.000
married	0.482	0.706	0.000
nodeg	0.373	0.301 v	0.000
re74	6214.351	13882.020	0.000
re75	6559.544	13513.232	0.000

Over lap indicator - for ATT

```
ind = ps>=min(ps[d1$treat==1]) & ps<=max(ps[d1$treat==1])
```

	Ave. Treat	Ave. control	T-test
age	26.921	28.730	0.000
ed	11.402	11.409	0.863
black	0.268	0.201	0.000
hisp	0.098	0.111	0.013
married	0.482	0.441	0.000
nodeg	0.373	0.447	0.000
re74	6214.351	4770.516	0.000
re75	6559.544	3833.817	0.000

IPSW: Imposing overlap, $0.1 < P - score < -0.9$

Over lap indicator - $0.1 < P - score < 0.9$

ind = $ps \geq 0.1$ & $ps \leq 0.9$

The second limitatin has no meaning as the max P-score is 0.48

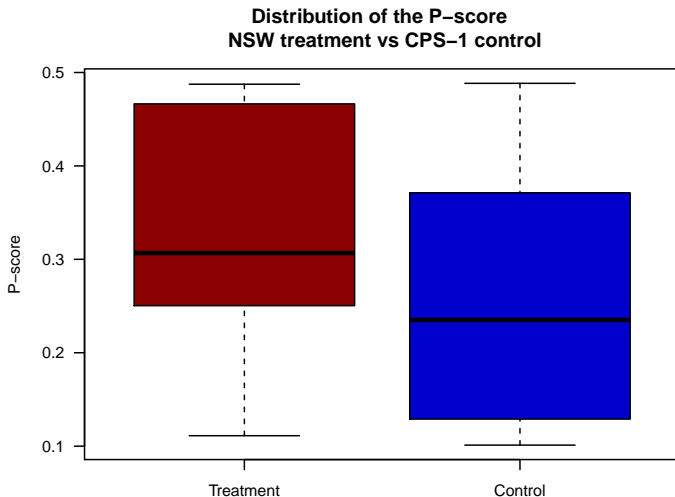
	Ave. Treat	Ave. control	T-test
age	26.485	27.316	0.176
ed	10.563	10.640	0.626
black	1.000	1.000	NA
hisp	0.000	0.000	NA
married	0.199	0.258	0.026
nodeg	0.631	0.599	0.300
re74	1972.791	1823.502	0.501
re75	1243.187	1081.554	0.153

IPSW: Imposing overlap, $0.1 < P - score < -0.9$

```
> summary(w[treat1==0 & ind])
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
1.113  1.148   1.307   1.390  1.590   1.955

> summary(w[treat1==1 & ind])
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
2.052  2.144   3.260   3.499  3.994   8.992
```

IPSW: Imposing overlap, $0.1 < P - \text{score} < -0.9$



IPSW: Imposing overlap, $0.1 < P - score < -0.9$

	Regression NSW	Regression CPS-1	IPSW	Matching on P-score
treat	1675.86** (639.34)	793.59 (548.25)	1249.49 (657.07)	985.86 (782.81)
age	141.73 (274.42)	-233.68*** (41.18)	357.59 (231.96)	234.64 (324.96)
ed	385.02 (230.11)	166.85*** (28.66)	236.52 (182.00)	368.07 (230.59)
R ²	0.06	0.48	0.05	0.22
Adj. R ²	0.03	0.48	0.03	0.22
Num. obs.	445	16177	504	9748

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

- What is your opinion on the IPSW results when limiting the data to $0.1 \leq P - score \leq 0.9$
- Is there a problem with this limitation on the data? **Yes**