

Section 10 : SATE to PATT

Andrew Bertoli

6 November 2013

Roadmap

1. Sandwich Estimator
2. From SATE to PATT
3. Reweighting when Matching

Sandwich Estimator

Motivation (for our purposes):

Often we want to use regression on experimental data to get a better estimate of the treatment effect.

Problem: Our standard errors will now be wrong since they are computed under the regression assumptions.

1. Randomization does not guarantee linear additivity
2. Randomization does not imply $\epsilon_i \sim N(0, \sigma^2)$ for all i .

Sandwich Estimator

To account for this problem, we estimate the standard errors using this formula

$$\hat{J}_n(\hat{\theta}_n)^{-1} \hat{V}_n(\hat{\theta}_n) \hat{J}_n(\hat{\theta}_n)^{-1}$$

where

$$\hat{J}_n(\hat{\theta}_n) = -E_g[l_n''(\theta)]$$

and

$$\hat{V}_n(\hat{\theta}_n) = \text{Var}_g(l_n'(\theta))$$

Sandwich Estimator

```
> library(sandwich)
```

```
b=bread(model)
```

```
m=meat(model)
```

```
s=sandwich(x=model,bread=b,meat=m)
```

Sandwich Estimator

> s

	(Intercept)	T	X2
(Intercept)	0.191727982	-0.0088418523	-0.0200045786
T	-0.008841852	0.0028719459	0.0008835502
X2	-0.020004579	0.0008835502	0.0021214117
X3	0.003083361	-0.0003325502	-0.0003981360

X3

(Intercept)	0.0030833607
T	-0.0003325502
X2	-0.0003981360
X3	0.0002719636

SATE to PATT

Motivation: Sometimes we have an experiment on a non-random sample of the population, and we want to estimate the treatment effect on all of the treated units in a population.

Example: A new dog food comes out. We want to test if the new dog food is causing dogs to gain weight.

SATE to PATT



SATE to PATT



Assumptions

1. Treatment means the same thing in both the randomized trial and the non-random study

2. SUTVA

3. Treatment Assignment Ignorability within the trial:

$$T \perp Y_{11}, Y_{10} | \mathbf{X}$$

4. Overlap between Trial Treated and Trial Control:

$$P(T = 1 | \mathbf{X} = \mathbf{x}, I = 1) < 1$$

5. Sample Assignment Ignorability for the treated unit experiment:

$$I \perp Y_{11}, Y_{01} | \mathbf{W}$$

6. Overlap between the Trial Treated and Population Treated:

$$0 < P(I = 1 | \mathbf{W} = \mathbf{w}, T = 1)$$

Notes: $I \in \{0, 1\}$ indicates whether the unit is in the experiment, and \mathbf{W} is the matrix of covariates used to reweight the treated units to resemble the population treated.

SATE to PATT

Question: How could these assumptions be violated in our study about dog obesity.

SATE to PATT

Potential Problems

1. Say we required the dogs in our study to eat 2 lbs of dog food per day, but not all the dogs in the population ate two pounds of dog food per day.
2. As the treated dogs become obese, the untreated dogs get less self-conscious and let themselves go.
- 3-4. The experiment was not really random.
5. There was some systematic difference between experimental dogs and the population treated dogs that we did not control for.
6. There were some population treated dogs that had no chance of being in the experiment.

SATE to PATT

Steps

1. Match the sample treated units to the sample control units, like you were trying to estimate the sample ATT.
2. Calculate weights for the sample treated units so that they resemble the population treated units.
3. Verify that the average outcome for the reweighted treated units is similar to the average outcome for the population treated units. If this is not true, something probably went wrong.
4. Estimate the ATT with these weights. This will give you the estimate of the PATT.

SATE to PATT

Data

We start with 100,000 dogs.

We know the height and weight for each dog.

680 dogs were selected into our experiment. Taller dogs were more likely to be in our experiment.

Some of the remaining dogs got the new food, and others did not. Heavier dogs were more likely to get the new food.

In reality, the new food increases each dog's weight by 10%.

SATE to PATT

Basic Terminology

Sample Treated: The treated dogs in our experiment

Sample Control: The control dogs in our experiment

Population Treated: The dogs that were not in our experiment whose owners decided to give them the new food

Population Control: The dogs that were not in our experiment whose owners decided not to give them the new food

SATE to PATT

```
> head(experiment)
```

	Outcome	I	Height	Weight	T
307	106.13599	1	47.88741	102.35123	0
323	31.62008	1	10.37685	23.52044	1
346	60.69045	1	27.58938	58.61883	0
387	51.88334	1	24.89269	48.35533	0
469	67.64650	1	33.75219	66.84402	1
675	19.22246	1	6.54694	19.32119	0

SATE to PATT

```
> head(population)
```

	Outcome	I	Height	Weight	T
1	86.33210	0	33.33105	71.79956	1
2	98.78965	0	49.87987	102.63939	0
3	104.12604	0	45.69805	96.98844	1
4	50.93354	0	22.40200	45.96683	0
5	57.25342	0	28.42934	58.31783	0
6	106.85577	0	49.35578	100.92963	1

SATE to PATT

Step 1. Match the sample treated units to the sample control units, like you wanted to estimate the sample ATT.

```
> library(Matching)
gen=GenMatch(T=experiment$T,X=cbind(experiment
$Weight,experiment$Height))

mat=Match(T=experiment$T, X=cbind(experiment
$Weight,experiment$Height), Weight.matrix=gen)

t.experiment=experiment[mat$index.treated,]
c.experiment=experiment[mat$index.control,]
```

SATE to PATT

Step 2. Calculate weights for the sample treated units so that they resemble the population treated units.

```
> pop.treated.means=apply(pop.treated[,3:4],2,mean)
eta=c(1,pop.treated.means)
constr=length(eta)
G=cbind(1,t.experiment[,3:4])
G=t(G)
lambda = rep(0, constr - 1)
q1 = rep(1 / length(lambda), ncol(G))
mt <- minxent.multiple(q=q1, G=G, eta=eta, lambda=lambda)
weights=c(mt$Estimates)
```

SATE to PATT

```
> wtd.t.test(t.experiment$Height, pop.treated  
$Height, weight=weights*300, samedata=FALSE)  
$test  
[1] "Two Sample Weighted T-Test (Welch)"
```

```
$coefficients
```

t.value	df	p.value
3.067026e-14	3.056040e+02	1.000000e+00

SATE to PATT

```
> wtd.t.test(t.experiment$Weight, pop.treated$Weight,  
weight=weights*300, samedata=FALSE)
```

```
$test
```

```
[1] "Two Sample Weighted T-Test (Welch)"
```

```
$coefficients
```

t.value	df	p.value
1.016571e-14	3.055904e+02	1.000000e+00

SATE to PATT

Step 3. Verify that the average outcome for the reweighted treated units is similar to the average outcome for the population treated units.

```
> wtd.t.test(t.experiment$Outcome, pop.treated$Outcome,  
weight=weights*300, samedata=FALSE)
```

```
$test
```

```
[1] "Two Sample Weighted T-Test (Welch)"
```

```
$coefficients
```

t.value	df	p.value
-0.2649108	305.7014039	0.7912567

```
$additional
```

Difference	Mean.x	Mean.y	Std. Err
-0.4062917	78.0956744	78.5019660	1.5336924

SATE to PATT

Step 4. Estimate the ATT with these weights. This will give you the estimate of the PATT.

```
> wtd.t.test(t.experiment$Outcome,c.experiment$Outcome,  
weight=weights*300,samedata=TRUE)
```

```
$test
```

```
[1] "Two Sample Weighted T-Test (Welch)"
```

```
$coefficients
```

	t.value	df	p.value
	3.329425e+00	5.919306e+02	9.243430e-04

```
$additional
```

Difference	Mean.x	Mean.y	Std. Err
6.843383	78.095674	71.252292	2.055425

SATE to PATT

Comparison of Estimated PATTs

Normal t-tests on experimental data: 6.1

Reweighted t-test on experimental data: 6.8

Actual PATT: 7.1

Reweighting when Matching

Say that we have an observational study with 1000 people who had the flu.

The outcome of interest is days until recovery.

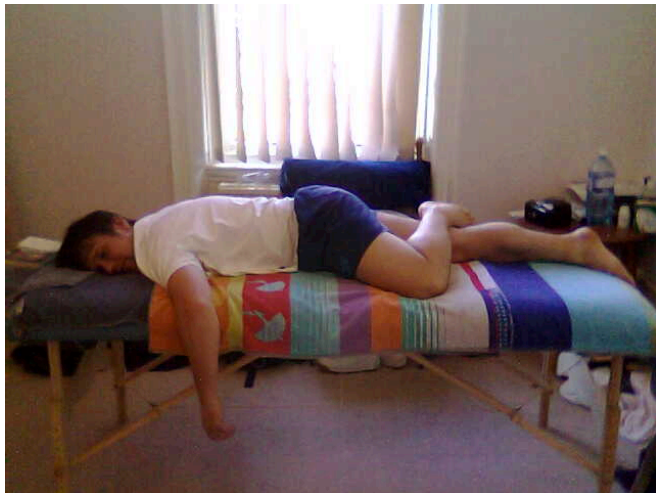
The two treatments are drinking orange juice and taking days off of work.

We want to know which is better.

Reweighting when Matching



Reweighting when Matching



Reweighting when Matching

One Approach

Step 1: Estimate the ATT for Orange Juice

Step 2: Estimate the ATT for Extra Rest

Step 3: Compare the Estimates

Even if we get very good estimates, what is the problem with this approach?

Answer: The people who took orange juice will be a different population than the people who got extra rest. Therefore, even if we know the true causal effect in both populations, they are defined on different populations and should not be compared directly.

Reweighting when Matching

Better Approach

Basic Idea: Estimate the effect of extra rest for people who drank orange juice.

Step 1. Find matches for all the people with who drank orange juice using all the people who got extra rest as controls.

Step 2. Find matches for all the people with who drank orange juice using all the people who did not get extra rest as controls.

Step 3. Use a t-test to compare the matches from Step 1 to the matches from Step 2.

Step 4. Compare this estimate to the estimated ATT of orange juice, calculated in the usual way.

Now we will be using the same population to compare the effects of the treatments.

Reweighting when Matching

```
> OJ=data[data$OrangeJuice==1,]
```

```
Sleepers=data[data$ExtraRest==1,]
```

```
NonSleepers=data[data$ExtraRest==0,]
```

Reweighting when Matching

```
> X=c(OJ$Age,Sleepers$Age)
```

```
T=c(rep(1, sum(data$OrangeJuice)),rep(0, sum(data  
$ExtraRest)))
```

```
mat1=Match(Tr=T,X=X,estimand="ATT")
```

Reweighting when Matching

```
> X=c(OJ$Age,NonSleepers$Age)
```

```
T=c(rep(1, sum(data$OrangeJuice)),rep(0, 1000-sum(data  
$ExtraRest)))
```

```
mat2=Match(Tr=T,X=X,estimand="ATT")
```


Reweighting when Matching

```
t=Sleepers[mat1$index.control-sum(data$OrangeJuice),]  
c=NonSleepers[mat2$index.control-sum(data$OrangeJuice),]  
t.test(t$DaysSick,c$DaysSick,paired=TRUE)
```

Reweighting when Matching

We could also do the opposite approach by estimating the effect of drinking orange juice for the people who got extra rest.

Step 1. Find matches for all the people who got extra rest using all the people who drank orange juice as controls.

Step 2. Find matches for all the people who got extra rest using all the people who did not drink orange juice as controls.

Step 3. Use a t-test to compare the matches from Step 1 to the matches from Step 2.

Step 4. Compare this estimate to the estimated ATT of extra rest, calculated in the usual way.

Reweighting when Matching

This method is useful if you are doing a matching study and you want to test if one treatment was more effective than another treatment.

This method does not involve extrapolation, so it is better than reweighing one group to look like the other. Reweighting units involves extrapolation.

Questions

1. Lectures
2. Readings
3. Homework