

Section 11 : Sensitivity Analysis

Andrew Bertoli

13 November 2013

Roadmap

1. Sensitivity Analysis
2. Bootstrapping
3. Homework

Sensitivity Analysis

Model of an Observational Study

Unit i 's probability of treatment is

$$\pi_i = P(T_i = 1)$$

and its probability of control is

$$1 - \pi_i = P(T_i = 0)$$

So we have

$$P(T_1 = t_1, \dots, T_n = t_n) = \prod_{i=1}^n \pi_i^{t_i} (1 - \pi_i)^{1-t_i}$$

Sensitivity Analysis

Example

Question: Does spending time at the beach improve your health?

$$\pi_i = P(\text{Person } i \text{ Goes to the Beach Regularly})$$

$$1 - \pi_i = P(\text{Person } i \text{ Does Not Got to the Beach Regularly})$$

Now say that we have information on everyone's age and gender.

Sensitivity Analysis



Sensitivity Analysis

Overt Bias

An observational study is free of hidden bias if every π_i (though unknown) only depends on the observed covariates

$$\mathbf{x}_i = (\text{Age}_i, \text{Gender}_i)$$

Define the probability that Person i will be a beach person as an unknown function of \mathbf{x}_i

$$\pi_i = \lambda(\mathbf{x}_i)$$

So the probability of treatment assignment is

$$P(T_1 = t_1, \dots, T_n = t_n) = \prod_{i=1}^n \lambda(\mathbf{x}_i)^{t_i} (1 - \lambda(\mathbf{x}_i))^{1-t_i}$$

As long as we match exactly on age and gender, our study will be valid. We do not need to know the function λ .

Sensitivity Analysis

Of course, we rarely believe that a study is free of hidden bias.

So how would our inferences change for different magnitudes of hidden bias?

There is hidden bias if two units have the same observed covariates but different probabilities of treatment assignment.

In other words

$x_a = x_b$, but $\pi_a \neq \pi_b$ for some a and b .

Sensitivity Analysis

So assume that each person who goes to the beach is matched with exactly one person who doesn't go to the beach.

The treatment odds for Person a and Person b (in the pair) are

$$O_a = \frac{\pi_a}{1-\pi_a} \text{ and } O_b = \frac{\pi_b}{1-\pi_b}.$$

The ratio of these odds, $\Gamma = O_a/O_b$, measures bias after matching.

We also have

$$P(T_a = 1 | T_a + T_b = 1) = \frac{\pi_a(1-\pi_b)}{\pi_a(1-\pi_b) + \pi_b(1-\pi_a)}$$

Sensitivity Analysis

The key parameter in Rosenbaum's sensitivity analysis is the treatment odds ratio Γ .

$$\Gamma = O_a / O_b = \frac{\pi_a / (1 - \pi_a)}{\pi_b / (1 - \pi_b)}$$

Substantively, this is the ratio of the odds of the treated unit going to the beach regularly and the odds of the matched control unit going to the beach regularly.

Sensitivity Analysis

Sensitivity analysis involves seeing if our results hold for different ranges of Γ . Let p_a and p_b be the probability of each unit in the matched pair being treated, conditional on exactly one being treated.

So if $\Gamma = 1$, the treatment and control unit within each pair has the same value of treatment assignment ($p_a, p_b = 0.5$).

If $1/2 \leq \Gamma \leq 2$, no unit can be more than twice as likely as its match to get treated ($0.33 \leq p_a, p_b \leq 0.66$).

If $1/3 \leq \Gamma \leq 3$, no unit can be more than three times as likely as its match to get treated ($0.25 \leq p_a, p_b \leq 0.75$).

(The quick formula for getting these bounds is

$$\frac{1}{\Gamma+1} \leq p_a, p_b \leq \frac{\Gamma}{\Gamma+1})$$

Sensitivity Analysis

Question: Imagine that we are doing our beach study and there was no treatment effect. If $1/2 \leq \Gamma \leq 2$, what would be the worst way the odds could go to make it look like there was a treatment effect.

Answer: For every pair, the healthier person was twice as likely to go to the beach regularly, or the less healthy person is twice as likely to go to the beach regularly.

Sensitivity Analysis

So say that these are our four matched pairs. Let's start with $\Gamma = 1$.

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	P(T1=1 T1+T2=1)	Min Gamma	Max Gamma
1	6	5	1	0	13	13	1	1	0.5	1	1
2	3	7	1	0	22	22	1	1	0.5	1	1
3	4	7	1	0	45	45	0	0	0.5	1	1
4	7	14	1	0	70	70	0	0	0.5	1	1

Sensitivity Analysis

So say that these are our four matched pairs. Let's start with $\Gamma = 1$.

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	P(T1=1 T1+T2=1)	Min Gamma	Max Gamma
1	6	5	1	0	13	13	1	1	0.5	1	1
2	3	7	1	0	22	22	1	1	0.5	1	1
3	4	7	1	0	45	45	0	0	0.5	1	1
4	7	14	1	0	70	70	0	0	0.5	1	1

The Wilcoxon Signed-Rank test statistic W is ?.

Sensitivity Analysis

Reminder: Wilcoxon Signed Rank Test

1. Drop pairs where the matches have the same outcome
2. Calculate the difference in outcomes within each pair
3. Rank the pairs (Smallest Difference=1)
4. Take the sum of the ranks where the treated unit had the higher outcome

Sensitivity Analysis

So say that these are our four matched pairs. Let's start with $\Gamma = 1$.

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	P(T1=1 T1+T2=1)	Min	Gamma	Max	Gamma
1	6	5	1	0	13	13	1	1	0.5		1		1
2	3	7	1	0	22	22	1	1	0.5		1		1
3	4	7	1	0	45	45	0	0	0.5		1		1
4	7	14	1	0	70	70	0	0	0.5		1		1

The Wilcoxon Signed-Rank test statistic W is 1.

Sensitivity Analysis

So say that these are our four matched pairs. Let's start with $\Gamma = 1$.

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	P(T1=1 T1+T2=1)	Min	Gamma	Max	Gamma
1	6	5	1	0	13	13	1	1	0.5		1		1
2	3	7	1	0	22	22	1	1	0.5		1		1
3	4	7	1	0	45	45	0	0	0.5		1		1
4	7	14	1	0	70	70	0	0	0.5		1		1

The Wilcoxon Signed-Rank test statistic W is 1.

Sum of the ranks = $1+2+3+4=10$

Sensitivity Analysis

So say that these are our four matched pairs. Let's start with $\Gamma = 1$.

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	$P(T1=1 T1+T2=1)$	Min	Gamma	Max	Gamma
1	6	5	1	0	13	13	1	1	0.5		1		1
2	3	7	1	0	22	22	1	1	0.5		1		1
3	4	7	1	0	45	45	0	0	0.5		1		1
4	7	14	1	0	70	70	0	0	0.5		1		1

The Wilcoxon Signed-Rank test statistic W is 1.

Sum of the ranks = $1+2+3+4=10$

p-value = $P(W \in \{0, 1, 9, 10\} | \text{Randomization})$

Sensitivity Analysis

So say that these are our four matched pairs. Let's start with $\Gamma = 1$.

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	$P(T1=1 T1+T2=1)$	Min	Gamma	Max	Gamma
1	6	5	1	0	13	13	1	1	0.5		1		1
2	3	7	1	0	22	22	1	1	0.5		1		1
3	4	7	1	0	45	45	0	0	0.5		1		1
4	7	14	1	0	70	70	0	0	0.5		1		1

The Wilcoxon Signed-Rank test statistic W is 1.

Sum of the ranks = $1+2+3+4=10$

p-value = $P(W \in \{0, 1, 9, 10\} | \text{Randomization})$

$$\text{p-value} = \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} = \frac{1}{4}$$

Sensitivity Analysis

In an experimental setting, we compute our p-values by assuming randomness and asking how rare our test statistic would be if there was no treatment effect.

Specifically, we look at the distribution of the test statistic under random assignment and no treatment effect.

If we believe that there is some degree of hidden bias, we cannot compute exact p-values. But we can find the p-value for the worst case scenario (given a maximum Γ).

Sensitivity Analysis

Let's look at the case where $\frac{1}{2} \leq \Gamma \leq 2$ (no one is more than twice as likely to go to the beach as their matched partner).

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	P(T1=1 T1+T2=1)	Min	Gamma	Max	Gamma
1	6	5	1	0	13	13	1	1	0.667	0.5		2	
2	3	7	1	0	22	22	1	1	0.333	0.5		2	
3	4	7	1	0	45	45	0	0	0.333	0.5		2	
4	7	14	1	0	70	70	0	0	0.333	0.5		2	

Sensitivity Analysis

Let's look at the case where $\frac{1}{2} \leq \Gamma \leq 2$ (no one is more than twice as likely to go to the beach as their matched partner).

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	P(T1=1 T1+T2=1)	Min	Gamma	Max	Gamma
1	6	5	1	0	13	13	1	1	0.667	0.5		2	
2	3	7	1	0	22	22	1	1	0.333	0.5		2	
3	4	7	1	0	45	45	0	0	0.333	0.5		2	
4	7	14	1	0	70	70	0	0	0.333	0.5		2	

The Wilcoxon Signed-Rank test statistic W is still 1.

Sensitivity Analysis

Let's look at the case where $\frac{1}{2} \leq \Gamma \leq 2$ (no one is more than twice as likely to go to the beach as their matched partner).

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	P(T1=1 T1+T2=1)	Min	Gamma	Max	Gamma
1	6	5	1	0	13	13	1	1	0.667	0.5		2	
2	3	7	1	0	22	22	1	1	0.333	0.5		2	
3	4	7	1	0	45	45	0	0	0.333	0.5		2	
4	7	14	1	0	70	70	0	0	0.333	0.5		2	

The Wilcoxon Signed-Rank test statistic W is still 1.

$$\text{Sum(Ranks)}=1+2+3+4=10$$

Sensitivity Analysis

Let's look at the case where $\frac{1}{2} \leq \Gamma \leq 2$ (no one is more than twice as likely to go to the beach as their matched partner).

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	P(T1=1 T1+T2=1)	Min	Gamma	Max	Gamma
1	6	5	1	0	13	13	1	1	0.667	0.5		2	
2	3	7	1	0	22	22	1	1	0.333	0.5		2	
3	4	7	1	0	45	45	0	0	0.333	0.5		2	
4	7	14	1	0	70	70	0	0	0.333	0.5		2	

The Wilcoxon Signed-Rank test statistic W is still 1.

$$\text{Sum(Ranks)}=1+2+3+4=10$$

$$\text{p-value} = P(W \in \{0, 1, 9, 10\} | \text{This Randomization})$$

Sensitivity Analysis

Let's look at the case where $\frac{1}{2} \leq \Gamma \leq 2$ (no one is more than twice as likely to go to the beach as their matched partner).

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	P(T1=1 T1+T2=1)	Min	Gamma	Max	Gamma
1	6	5	1	0	13	13	1	1	0.667	0.5		2	
2	3	7	1	0	22	22	1	1	0.333	0.5		2	
3	4	7	1	0	45	45	0	0	0.333	0.5		2	
4	7	14	1	0	70	70	0	0	0.333	0.5		2	

The Wilcoxon Signed-Rank test statistic W is still 1.

$$\text{Sum(Ranks)}=1+2+3+4=10$$

$$\text{p-value} = P(W \in \{0, 1, 9, 10\} | \text{This Randomization})$$

$$\text{p-value} = \left(\frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3}\right) + \left(\frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3}\right) + \left(\frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3}\right) + \left(\frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3}\right)$$

Sensitivity Analysis

Let's look at the case where $\frac{1}{2} \leq \Gamma \leq 2$ (no one is more than twice as likely to go to the beach as their matched partner).

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	P(T1=1 T1+T2=1)	Min	Gamma	Max	Gamma
1	6	5	1	0	13	13	1	1	0.667	0.5		2	
2	3	7	1	0	22	22	1	1	0.333	0.5		2	
3	4	7	1	0	45	45	0	0	0.333	0.5		2	
4	7	14	1	0	70	70	0	0	0.333	0.5		2	

The Wilcoxon Signed-Rank test statistic W is still 1.

$$\text{Sum(Ranks)} = 1 + 2 + 3 + 4 = 10$$

$$\text{p-value} = P(W \in \{0, 1, 9, 10\} | \text{This Randomization})$$

$$\text{p-value} = \left(\frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3}\right) + \left(\frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3}\right) + \left(\frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3}\right) + \left(\frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3}\right)$$

$$\text{p-value} = \frac{3}{9}$$

Sensitivity Analysis

So we can say that our maximum p-value is $3/9$ as long as no person was more than twice as likely to go to the beach as their matched control.

Sensitivity Analysis

What if $\frac{1}{4} \leq \Gamma \leq 4$ (no one is more than twice as likely to go to the beach as their matched partner).

> data

	Y1	Y2	T1	T2	Age1	Age2	Gender1	Gender2	P(T1=1 T1+T2=1)	Min	Gamma	Max	Gamma
1	6	5	1	0	13	13	1	1	0.8	0.25		4	
2	3	7	1	0	22	22	1	1	0.2	0.25		4	
3	4	7	1	0	45	45	0	0	0.2	0.25		4	
4	7	14	1	0	70	70	0	0	0.2	0.25		4	

The Wilcoxon Signed-Rank test statistic W is still 1.

$$\text{Sum(Ranks)}=1+2+3+4=10$$

$$\text{p-value} = P(W \in \{0, 1, 9, 10\} | \text{This Randomization})$$

$$\text{p-value} = \left(\frac{1}{5} \cdot \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{1}{5}\right) + \left(\frac{4}{5} \cdot \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{1}{5}\right) + \left(\frac{1}{5} \cdot \frac{4}{5} \cdot \frac{4}{5} \cdot \frac{4}{5}\right) + \left(\frac{4}{5} \cdot \frac{4}{5} \cdot \frac{4}{5} \cdot \frac{4}{5}\right)$$

$$\text{p-value} = \frac{3}{9}$$

Bootstrapping (from Section 5)

Motivation

Imagine that we want to study the height of people in the United States, so we take a random sample of 10,000 Americans and measure them.

By the Central Limit Theorem, the sample average $\hat{\mu}$ will be distributed about $N(\mu, \sigma/100)$, where μ is the population mean and σ is the sample variance.

However, what if we are interested in the population median? The CLT does not give us a simple way to estimate its distribution.

Bootstrapping (from Section 5)

Basic Idea

To estimate the distribution of a parameter like the sample median

1. Take many random samples (with replacement) from the original sample, each of size n (in this case 10,000)
2. Estimate the parameter for each sample
3. The distribution of these estimates should be close to the sampling distribution of the estimator

Bootstrapping (from Section 5)

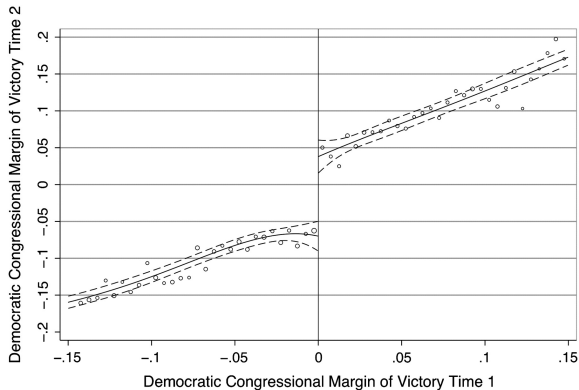
Basic Idea

The previous example is a case of non-parametric bootstrapping, because we made no assumptions about the distribution of the heights of Americans. We simply had a random sample of the population, and we resampled from our sample over and over.

1. Take many random samples (with replacement) from the original sample, each of size n (in this case 10,000)
2. Estimate the parameter for each sample
3. The distribution of these estimates should be close to the sampling distribution of the estimator

Regression Discontinuity (from Section 5)

Creating the Graph



Regression Discontinuity (from Section 5)

Estimating the Confidence Interval

1. The 95% confidence interval is found by bootstrapping
2. Starting with the left side of the cut-point, randomly sample (with replacement) from the points on the left. The size of the new sample should be the same as the number of points on the left.
3. Construct a regression line for this new sample
4. Repeat this process several thousand times. This will result in thousands of regression lines. Each regression line will be defined by a large number of (x, y) coordinates. Thus, each x will be associated with a large number of y 's.
5. For each x , find the lower 2.5% and upper 97.5% quantile for the y 's. These points will define the lower and upper bounds of the 95% confidence interval.

Questions

1. Lectures
2. Readings
3. Homework