Lecture 14

Annoucements

- Quiz 6
- Midterm is on 03/18/2020 (Wednesday)

Outline

- Randomized Control Trial (RCT)
  - Balance test
  - Result table
  - Attrition

Randomized Control Trial (RCT)

- Randomized control trial is also called random assignment or simply randomization or experiment
- They really mean the same thing: randomly assign people to treatment and control (comparison) group
- It is the best method to eliminate OVB/selection bias
  - If health insurance is randomly assigned, then People in treatment and control groups are not only similar in observables but also in unobservables
  - We are comparing apples to apples-No selection bias.
- The downside: too costly; ethical concerns.

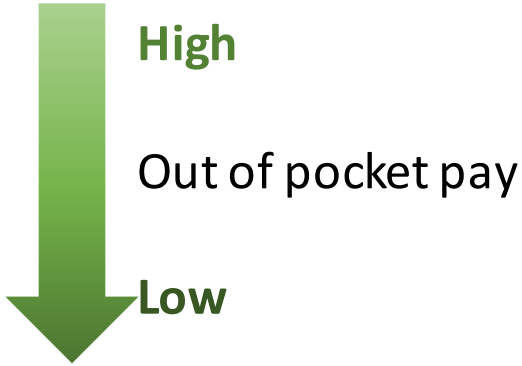
Example: RAND Health Insurance Experiment (HIE)

- Is it a good thing to make health insurance free?
- What unintended consequences might happen?
  - Cons: Moral hazard (cost)
  - Pros: Improved health? (benefit)
- RAND HIE
  - Families were **randomly assigned** to one of four health insurance plans
  - Don't pay premium, but out-of-pocket copay varies (you may think these different plans have different prices)
- Objective:
  - How does cost-sharing affect usage of health care?
  - Does increased usage of healthcare lead to better health outcomes?

Treatment and control group

- Control Group: Catastrophic Plan-(mimic no insurance)

- Treatment Group 1: Deductible
- Treatment Group 2: Coinsurance
- Treatment Group 3: Free



How do we evaluate randomized control trial?

- Is this truly a randomized experiment?
  - The meaning of random: Treatment and control group should look similar in terms of demographic characteristics and outcome variable before the experiment
  - How do you test it?
    - Show a table containing the means of demographic variables for treatment and control group before the experiment
    - Test if their difference is statistically different from 0
    - This is called a **balance test (balance table)**
    - The function of balance table is to test whether there is omitted variable/selection bias problem before the experiment

RCT: Balance table and Result Table

- Balance table: to test whether **demographic variables** are similar in treatment and control groups **before the experiment**

- Result table: to see if there is any difference in the **outcome variable after the experiment**

Exercise:

- Is this a balance table or result table?
- Is the age difference between coinsurance and catastrophic (control group) economically big?
- Is the difference statistically significant?

	Means	Differences between plan groups			
	Catastrophic plan (1)	Deductible – catastrophic (2)	Coinsurance – catastrophic (3)	Free – catastrophic (4)	Any insurance – catastrophic (5)
A. Demographic characteristics					
Female	.560	−.023	−.025	−.038	−.030
		(.016)	(.015)	(.015)	(.013)
Nonwhite	.172	−.019	−.027	−.028	−.025
		(.027)	(.025)	(.025)	(.022)
Age	32.4	.56	.97	.43	.64
	[12.9]	(.68)	(.65)	(.61)	(.54)
Education	12.1	−.16	−.06	−.26	−.17
	[2.9]	(.19)	(.19)	(.18)	(.16)
Family income	31,603	−2,104	970	−976	−654
	[18,148]	(1,384)	(1,389)	(1,345)	(1,181)
Hospitalized last year	.115	.004	−.002	.001	.001
		(.016)	(.015)	(.015)	(.013)
B. Baseline health variables					
General health index	70.9	−1.44	.21	−1.31	−.93
	[14.9]	(.95)	(.92)	(.87)	(.77)
Cholesterol (mg/dl)	207	−1.42	−1.93	−5.25	−3.19
	[40]	(2.99)	(2.76)	(2.70)	(2.29)
Systolic blood pressure (mm Hg)	122	2.32	.91	1.12	1.39
	[17]	(1.15)	(1.08)	(1.01)	(.90)
Mental health index	73.8	−.12	1.19	.89	.71
	[14.3]	(.82)	(.81)	(.77)	(.68)
Number enrolled	759	881	1,022	1,295	3,198

Before the experiment:
"Balance test"-a test for randomization

-0.023: ($Mean_{treat}$-$Mean_{control}$)=difference in proportion of female between treatment and control

0.016:standard error

Example: conduct hypothesis testing at α=5%

H_0: diff=0

t*=(-0.023-0)/0.016=1.43
$t_{\alpha/2}$ for n=500 and α=5% is 1.97
t*< $t_{\alpha/2}$. Fail to reject H_0.

Conclusion: Treatment and control groups have similar proportion of females

Randomization is good (for this variable)!

	Means	Differences between plan groups			
	Catastrophic plan (1)	Deductible – catastrophic (2)	Coinsurance – catastrophic (3)	Free – catastrophic (4)	Any insurance – catastrophic (5)
A. Demographic characteristics					
Female	.560	−.023	−.025	−.038	−.030
		(.016)	(.015)	(.015)	(.013)
Nonwhite	.172	−.019	−.027	−.028	−.025
		(.027)	(.025)	(.025)	(.022)
Age	32.4	.56	.97	.43	.64
	[12.9]	(.68)	(.65)	(.61)	(.54)
Education	12.1	−.16	−.06	−.26	−.17
	[2.9]	(.19)	(.19)	(.18)	(.16)
Family income	31,603	−2,104	970	−976	−654
	[18,148]	(1,384)	(1,389)	(1,345)	(1,181)
Hospitalized last year	.115	.004	−.002	.001	.001
		(.016)	(.015)	(.015)	(.013)
B. Baseline health variables					
General health index	70.9	−1.44	.21	−1.31	−.93
	[14.9]	(.95)	(.92)	(.87)	(.77)
Cholesterol (mg/dl)	207	−1.42	−1.93	−5.25	−3.19
	[40]	(2.99)	(2.76)	(2.70)	(2.29)
Systolic blood pressure (mm Hg)	122	2.32	.91	1.12	1.39
	[17]	(1.15)	(1.08)	(1.01)	(.90)
Mental health index	73.8	−.12	1.19	.89	.71
	[14.3]	(.82)	(.81)	(.77)	(.68)
Number enrolled	759	881	1,022	1,295	3,198

Exercise

1. Which group is the control group? How many treatment groups do we have?
2. Look at the coefficient in the red circle. What does it mean?
3. What is the standard error for that coefficient?

	Means	Differences between plan groups			
	Catastrophic plan (1)	Deductible – catastrophic (2)	Coinsurance – catastrophic (3)	Free – catastrophic (4)	Any insurance – catastrophic (5)
A. Demographic characteristics					
Female	.560	−.023	−.025	−.038	−.030
		(.016)	(.015)	(.015)	(.013)
Nonwhite	.172	−.019	−.027	−.028	−.025
		(.027)	(.025)	(.025)	(.022)
Age	32.4	.56	.97	.43	.64
	[12.9]	(.68)	(.65)	(.61)	(.54)
Education	12.1	−.16	−.06	−.26	−.17
	[2.9]	(.19)	(.19)	(.18)	(.16)

4. To conduct statistical inference for that coefficient. What is the null hypothesis? What is the t-stat?
5. Use d.f.=500 and a=0.05 to conclude whether you reject null hypothesis
6. **What does reject the null means in this example?**
7. **Is the result what you would like to see (the purpose of randomization)?**

Results of RCT

- If the randomization is successful, a simple mean comparison (naïve comparison) will provide the causal impact of treatment on the output
- In regression framework, the naive mean comparison is provided by simple regression:
- Health=β_0+β_1Insured+ε
- $\widehat{\beta_1}$ are our main results. We have different measures for health, so we end up with different $\widehat{\beta_1}$

- What are the two main sets of results?
- Does giving people free insurance increase medical care usage?
- Does it lead to better health?

	Means	Differences between plan groups			
	Catastrophic plan (1)	Deductible – catastrophic (2)	Coinsurance – catastrophic (3)	Free – catastrophic (4)	Any insurance – catastrophic (5)
A. Health-care use					
Face-to-face visits	2.78 [5.50]	.19 (.25)	.48 (.24)	1.66 (.25)	.90 (.20)
Outpatient expenses	248 [488]	42 (21)	60 (21)	169 (20)	101 (17)
Hospital admissions	.099 [.379]	.016 (.011)	.002 (.011)	.029 (.010)	.017 (.009)
Inpatient expenses	388 [2,308]	72 (69)	93 (73)	116 (60)	97 (53)
Total expenses	636 [2,535]	114 (79)	152 (85)	285 (72)	198 (63)
B. Health outcomes					
General health index	68.5 [15.9]	−.87 (.96)	.61 (.90)	−.78 (.87)	−.36 (.77)
Cholesterol (mg/dl)	203 [42]	.69 (2.57)	−2.31 (2.47)	−1.83 (2.39)	−1.32 (2.08)
Systolic blood pressure (mm Hg)	122 [19]	1.17 (1.06)	−1.39 (.99)	−.52 (.93)	−.36 (.85)
Mental health index	75.5 [14.8]	.45 (.91)	1.07 (.87)	.43 (.83)	.64 (.75)
Number enrolled	759	881	1,022	1,295	3,198

After the experiment:
Result of RCT

There are two types of main results

- usage of medical care
- health status

What does the coefficient 0.19 mean?

- The average face-to-face visits for the deductible group is 0.19 unit higher than the average of the control group (catastrophic plan)

- Can you conduct statistical inference based on its standard error? Use d.f.=500 and α=0.05

	Means	Differences between plan groups			
	Catastrophic plan (1)	Deductible – catastrophic (2)	Coinsurance – catastrophic (3)	Free – catastrophic (4)	Any insurance – catastrophic (5)
		A. Health-care use			
Face-to-face visits	2.78 [5.50]	.19 (.25)	.48 (.24)	1.66 (.25)	.90 (.20)
Outpatient expenses	248 [488]	42 (21)	60 (21)	169 (20)	101 (17)
Hospital admissions	.099 [.379]	.016 (.011)	.002 (.011)	.029 (.010)	.017 (.009)
Inpatient expenses	388 [2,308]	72 (69)	93 (73)	116 (60)	97 (53)
Total expenses	636 [2,535]	114 (79)	152 (85)	285 (72)	198 (63)
		B. Health outcomes			
General health index	68.5 [15.9]	−.87 (.96)	.61 (.90)	−.78 (.87)	−.36 (.77)
Cholesterol (mg/dl)	203 [42]	.69 (2.57)	−2.31 (2.47)	−1.83 (2.39)	−1.32 (2.08)
Systolic blood pressure (mm Hg)	122 [19]	1.17 (1.06)	−1.39 (.99)	−.52 (.93)	−.36 (.85)
Mental health index	75.5 [14.8]	.45 (.91)	1.07 (.87)	.43 (.83)	.64 (.75)
Number enrolled	759	881	1,022	1,295	3,198

Exercise

1. Explain what the coefficient in the red circle means
2. Conduct statistical inference for this coef.
3. What is your conclusion? Did you identify a statistically significant difference between the treatment and control?

	Means	Differences between plan groups			
	Catastrophic plan (1)	Deductible – catastrophic (2)	Coinsurance – catastrophic (3)	Free – catastrophic (4)	Any insurance – catastrophic (5)
		A. Health-care use			
Face-to-face visits	2.78 [5.50]	.19 (.25)	.48 (.24)	1.66 (.25)	.90 (.20)
Outpatient expenses	248 [488]	42 (21)	60 (21)	169 (20)	101 (17)
Hospital admissions	.099 [.379]	.016 (.011)	.002 (.011)	.029 (.010)	.017 (.009)
Inpatient expenses	388 [2,308]	72 (69)	93 (73)	116 (60)	97 (53)
Total expenses	636 [2,535]	114 (79)	152 (85)	285 (72)	198 (63)

Evaluate RCT-attrition problem

- Attrition: participants drop out of the program before experiment is finished
- Attrition is a problem only when attribution is not balanced across the treatment and control group
- It is not a problem if attribution is random
- e.g. If more poor people in the **control** group drop out of the experiment than treatment group, health status for control group after the experiment will appear to be higher than it is without attrition (Because poor people tend to have poor health, and these poor people drop out in the control)
- **Treatment effect=health of treatment-health of control** (after the experiment)
  - Health of control increases, therefore treatment effect decreases
- Downward bias the effect by inflating health status of the control group

Exercise:

- In the previous example, what happens to the treatment effect if more educated (with a college degree) people in the treatment group drop out than the control? (Maybe they have better choices than free insurance). Will this attribution problem cause biased estimate? Upward or downward bias? Explain.

Review for quizzes

- How to read balance table and result table of RCT
- Be able to perform statistical inference for the balance table and result table
- Understand the attrition problem in RCT