Resultados Potenciales

2025-04-09

Modelo de Resultados Potenciales

Plan

- Contexto
- Notación de Resultados Potenciales
- Descomposición de la Diferencia Simple
- Supuesto de Independencia
- Supuesto SUTVA

Causalidad

- Objetivo 1: entender la notación de resultados potenciales
- Objetivo 2: poder calcular distintos conceptos de resultados potenciales
- Objetivo 3: entender la descomposición de la diferencia simple

Repaso

En la clase pasada aprendimos sobre diagramas causales.

Una herramienta muy útil para modelar el proceso generador de datos y entender qué variación es "buena" o "mala", y qué debe mantenerse constante.

Algunas limitaciones:

- Puede ser difícil trabajar con modelos complejos (en Economía no suele ser gran problema).
- No sirve bien para sistemas con retroalimentación. Especialmente complicado para sistemas de oferta y demanda.
- Es simple, pero pierde precisión.

Resultados Potenciales

El modelo de resultados potenciales es otra herramienta para pensar la causalidad.

Los DAGs son relativamente nuevos y tienen sus usos, pero los resultados potenciales son fundamentales.

La principal diferencia con la econometría tradicional es un tratamiento explícito de la causalidad y más énfasis en efectos heterogéneos.

Resultados Potenciales

- Pensemos en un “tratamiento”

Resultados Potenciales

- Pensemos en un "tratamiento"
  - Una nueva medicina, un programa de gobierno, exposición a una campaña publicitaria

Resultados Potenciales

- Pensemos en un "tratamiento"
  - Una nueva medicina, un programa de gobierno, exposición a una campaña publicitaria
- Para cada persona hay dos futuros posibles:

Resultados Potenciales

- Pensemos en un “tratamiento”
  - Una nueva medicina, un programa de gobierno, exposición a una campaña publicitaria
- Para cada persona hay dos futuros posibles:
  - uno donde recibe el tratamiento

Resultados Potenciales

- Pensemos en un "tratamiento"
  - Una nueva medicina, un programa de gobierno, exposición a una campaña publicitaria
- Para cada persona hay dos futuros posibles:
  - uno donde recibe el tratamiento
  - uno donde no lo recibe

Resultados Potenciales

- Pensemos en un "tratamiento"
  - Una nueva medicina, un programa de gobierno, exposición a una campaña publicitaria
- Para cada persona hay dos futuros posibles:
  - uno donde recibe el tratamiento
  - uno donde no lo recibe
- Nos interesa cómo el tratamiento afecta alguna variable de interés, el resultado.

Resultados Potenciales

- El **efecto causal** para cada individuo es la diferencia entre su resultado en ambos mundos.

Resultados Potenciales

- El **efecto causal** para cada individuo es la diferencia entre su resultado en ambos mundos.
- ¡Fácil! ¡Solo resta!

Resultados Potenciales

- El **efecto causal** para cada individuo es la diferencia entre su resultado en ambos mundos.
- ¡Fácil! ¡Solo resta!
- Pero: ¡solo observamos un resultado!

Resultados Potenciales

- El **efecto causal** para cada individuo es la diferencia entre su resultado en ambos mundos.
- ¡Fácil! ¡Solo resta!
- Pero: ¡solo observamos un resultado!
- El problema fundamental de la inferencia causal: datos faltantes

Notación

- Y representa los resultados
- El subíndice i denota un individuo
  - Puede ser una persona, empresa, estado, familia
- El superíndice 0 o 1 representa en qué universo estamos

Por lo tanto:

- Y_i^0 es el resultado si i *no recibe* el tratamiento
- Y_j^1 es el resultado si j *recibe* el tratamiento

Notación

- Sin superíndice significa el resultado observado
  - Y^0_{Joao} es el resultado para Joao, sin tratamiento
  - Y^1_{Joao} es el resultado para Joao, con tratamiento
  - Y_{Joao} es el resultado observado para Joao

Notación

- D_i representa el estado de tratamiento: 1 si tratado, 0 si no tratado
- Con esta notación:

$$Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0$$

- Efecto del tratamiento:

$$\delta_i = Y_i^1 - Y_i^0$$

Ejemplo

Individuo	Y_i^0	Y_i^1	δ_i
Ana	5	4	
Bruno	3	9	
Carla	3	2	
Daniel	7	4	

Ejemplo

Individuo	Y_i^0	Y_i^1	δ_i
Ana	5	4	$4 - 5 = -1$
Bruno	3	9	$9 - 3 = +6$
Carla	3	2	$2 - 3 = -1$
Daniel	7	4	$4 - 7 = -3$

Ejemplo

Lamentablemente, no podemos calcular δ_i, porque solo tenemos o Y_i^0 o Y_i^1.

Individuo	Y_i^0	Y_i^1	D_i	δ_i
Ana	?	4	1	$4-? =?$
Bruno	?	9	1	$9-? =?$
Carla	3	?	0	$?-3=?$
Daniel	7	?	0	$?-7=?$

Efectos Promedio del Tratamiento

No podemos calcular el efecto individual, pero *en algunos casos* podemos calcular *efectos promedio*.

Como cada persona tiene su propio efecto causal, no podemos hablar de "el efecto causal". Hay que ser más precisos.

El primer efecto que nos interesa es el *Efecto Promedio del Tratamiento* (ATE):

$$E[\delta_i] = E[Y_i^1 - Y_i^0] = E[Y_i^1] - E[Y_i^0]$$

En palabras, el ATE es simplemente el promedio de todos los efectos.

Efectos Promedio del Tratamiento

También podemos calcular efectos promedio para un grupo específico. Por ejemplo, si nos interesan las personas con cierta característica $X = x$:

$$E[\delta_i|X_i = x] = E[Y_i^1 - Y_i^0|X_i = x] = E[Y_i^1|X_i = x] - E[Y_i^0|X_i = x]$$

X puede ser género, ingreso, empleo, etc.

Efectos Promedio del Tratamiento

Una característica que suele importarnos: haber recibido tratamiento.

Efecto Promedio del Tratamiento para los Tratados (ATT):

$$E[\delta_i | D_i = 1] = E[Y_i^1 - Y_i^0 | D_i = 1] = E[Y_i^1 | D_i = 1] - E[Y_i^0 | D_i = 1]$$

Efectos Promedio del Tratamiento

Individuo	Y_i^0	Y_i^1	D_i	δ_i
Ana	5	4	1	-1
Bruno	3	9	1	$+6$
Carla	3	2	0	-1
Daniel	7	4	0	-3

¿Cuál es el ATE?

Efectos Promedio del Tratamiento

Individuo	Y_i^0	Y_i^1	D_i	δ_i
Ana	5	4	1	-1
Bruno	3	9	1	$+6$
Carla	3	2	0	-1
Daniel	7	4	0	-3

¿Qué pasa si calculamos el efecto promedio para mujeres?

Efectos Promedio del Tratamiento

Individuo	Y_i^0	Y_i^1	D_i	δ_i
Ana	5	4	1	-1
Bruno	3	9	1	$+6$
Carla	3	2	0	-1
Daniel	7	4	0	-3

¿Qué pasa con el ATT?

Ejemplo

Imaginemos un nuevo procedimiento médico.

El tratamiento es este nuevo procedimiento, y el control es seguir el procedimiento tradicional.

El resultado es la cantidad de años de vida del paciente bajo cada procedimiento.

Ejemplo

Paciente	Y_i^1	Y_i^0
1	7	1
2	5	6
3	5	1
4	7	8
5	4	2
6	10	1
7	1	10
8	5	6
9	3	7
10	9	8

Ejemplo

Paciente	Y_i^1	Y_i^0	δ_i
1	7	1	?
2	5	6	?
3	5	1	?
4	7	8	?
5	4	2	?
6	10	1	?
7	1	10	?
8	5	6	?
9	3	7	?
10	9	8	?

Calculemos los efectos del tratamiento para cada paciente.

Ejemplo

Paciente	Y_i^1	Y_i^0	δ_i
1	7	1	6
2	5	6	-1
3	5	1	4
4	7	8	-1
5	4	2	2
6	10	1	9
7	1	10	-9
8	5	6	-1
9	3	7	-4
10	9	8	1

- ¿Cuál es el ATE?

Ejemplo

$E[\delta_i] = \frac{1}{10} \sum_i \delta_i = 0.6$

Ejemplo

Paciente	Y_i^1	Y_i^0	δ_i	Sexo
1	7	1	6	H
2	5	6	-1	M
3	5	1	4	H
4	7	8	-1	M
5	4	2	2	H
6	10	1	9	M
7	1	10	-9	H
8	5	6	-1	M
9	3	7	-4	H
10	9	8	1	M

- ¿Cuál es el ATE para hombres? ¿Y para mujeres?

Ejemplo

- Hombres: -0.1
- Mujeres: $+0.7$

Ejemplo

Imaginemos un médico perfecto que conoce los resultados potenciales de cada paciente.

Prescribe el tratamiento si el efecto es positivo.

Ejemplo

Paciente	Y_i^1	Y_i^0	δ_i	D_i
1	7	1	6	1
2	5	6	-1	0
3	5	1	4	1
4	7	8	-1	0
5	4	2	2	1
6	10	1	9	1
7	1	10	-9	0
8	5	6	-1	0
9	3	7	-4	0
10	9	8	1	1

- ¿Cuál es el ATT? ¿Y el ATU?

Ejemplo

- ATT: $E[\delta_i|D_i = 1] = \frac{1}{5}\sum_{D_i=1}\delta_i = 4.4$
- ATU: $E[\delta_i|D_i = 0] = \frac{1}{5}\sum_{D_i=0}\delta_i = -3.2$

Ejemplo

Pero en la realidad, no observamos todos los resultados potenciales.

Solo vemos el resultado realizado.

Ejemplo

Paciente	Y_i^1	Y_i^0	δ_i	D_i	Y_i
1	7	1	6	1	7
2	5	6	-1	0	6
3	5	1	4	1	5
4	7	8	-1	0	8
5	4	2	2	1	4
6	10	1	9	1	10
7	1	10	-9	0	10
8	5	6	-1	0	6
9	3	7	-4	0	7
10	9	8	1	1	9

Ejemplo

Paciente	D_i	Y_i
1	1	7
2	0	6
3	1	5
4	0	8
5	1	4
6	1	10
7	0	10
8	0	6
9	0	7
10	1	9

¿Qué pasa si calculamos la diferencia simple?

Ejemplo

$$\frac{1}{5}\sum_{D_i=1} Y_i - \frac{1}{5}\sum_{D_i=0} Y_i = 7 - 7.4 = -0.4$$

Sesgo por Diferencia Simple

- Usando la diferencia simple, encontramos un efecto negativo ¡aunque el efecto promedio verdadero es positivo!
- Vamos a entender qué está pasando.

Sesgo por Diferencia Simple

Partimos de la diferencia simple:

$$\mathbb{E}[Y \mid D = 1] - \mathbb{E}[Y \mid D = 0] = \mathbb{E}[Y^1 \mid D = 1] - \mathbb{E}[Y^0 \mid D = 0]$$

Sesgo por Diferencia Simple

Partimos de la diferencia simple:

$$\mathbb{E}[Y \mid D = 1] - \mathbb{E}[Y \mid D = 0] = \mathbb{E}[Y^1 \mid D = 1] - \mathbb{E}[Y^0 \mid D = 0]$$

Luego sumamos y restamos $\mathbb{E}[Y^0 \mid D = 1]$:

$$\begin{aligned}\mathbb{E}[Y^1 \mid D = 1] - \mathbb{E}[Y^0 \mid D = 0] = \mathbb{E}[Y^1 \mid D = 1] {\color{red}- \mathbb{E}[Y^0 \mid D = 1]} + \\ {\color{blue}\mathbb{E}[Y^0 \mid D = 1]} - \mathbb{E}[Y^0 \mid D = 0]\end{aligned}$$

Expresión Final

$$\mathbb{E}[Y \mid D=1] - \mathbb{E}[Y \mid D=0] =$$

$$\underbrace{\mathbb{E}[Y^1 \mid D=1] - \mathbb{E}[Y^0 \mid D=1]}_{\text{ATT}} + \underbrace{\mathbb{E}[Y^0 \mid D=1] - \mathbb{E}[Y^0 \mid D=0]}_{\text{Sesgo de Selección}}$$

Sesgo de Selección

En el ejemplo del médico perfecto, ATT $= 4.4$ y diferencia simple $=$ -0.4.

Vamos a calcular: $\mathbb{E}[Y^0 \mid D = 1] - \mathbb{E}[Y^0 \mid D = 0]$

Sesgo de Selección

(13 - 37)/5 = -24/5 = -4.8

$$\underbrace{\text{Diferencia Simple}}_{-0.4} = \underbrace{ATT}_{4.4} + \underbrace{\text{Sesgo de Selección}}_{-4.8}$$

Descomposición de la Diferencia Simple

- Interpretación: El sesgo de selección mide cuán distintos habrían sido los tratados y no tratados *si no hubieran recibido tratamiento*.
- En el ejemplo, el sesgo es negativo porque los pacientes que habrían salido peor sin tratamiento fueron los que sí lo recibieron.

Supuesto de Independencia

- Introducimos un supuesto para corregir el problema.
- Supuesto de independencia:

$$(Y^1, Y^0) \perp D$$

- También se llama asignación aleatoria del tratamiento
- Significa que ser tratado o no no depende de los resultados potenciales

Supuesto de Independencia

Recordemos: si $A \perp B$, entonces $P(A|B) = P(A)$.

Entonces, el supuesto de independencia nos da:

- El sesgo de selección es cero:

$$\text{Sesgo} = E[Y^0|D=1] - E[Y^0|D=0] = E[Y^0] - E[Y^0] = 0$$

Supuesto de Independencia

- Segundo, la diferencia simple identifica el ATT:

$$
\begin{aligned}
ATT =& E[Y|D=1] - E[Y|D=0] \\
=& E[Y^1|D=1] - E[Y^0|D=1] \\
=& E[Y^1] - E[Y^0]
\end{aligned}
$$

Supuesto de Independencia

- Tercero, el ATE es igual al ATT:

$$ATE = E[Y^1] - E[Y^0] = ATT$$

Supuesto de Independencia

Bajo asignación aleatoria:

- No hay sesgo de selección
- La diferencia simple estima el ATT
- ATT = ATE

SUTVA

Asumimos también SUTVA: *Asignación Estable del Valor del Tratamiento por Unidad*.

SUTVA

Este supuesto excluye tres cosas:

- Dosis distintas. El tratamiento es 0 o 1. Nada de medias dosis ni calidades distintas.
- Efectos indirectos: el tratamiento de uno no afecta el resultado de otro.
- Efectos de equilibrio general.

Resultados de Simulación

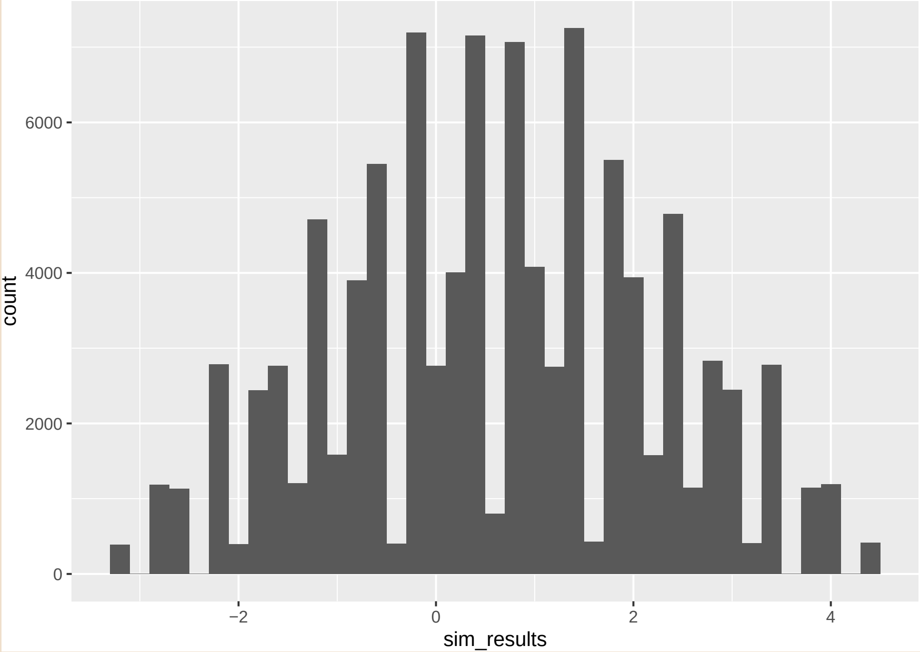
```r
y1 <- c(7,5,5,7,4,10,1,5,3,9)
y0 <- c(1,6,1,8,2,1,10,6,7,8)

simple_diff <- function() {
  treated_indices <- sample(10, 5)
  untreated_indices <- setdiff(1:10, treated_indices)
  treated_mean <- mean(y1[treated_indices])
  untreated_mean <- mean(y0[treated_indices])
  return(treated_mean - untreated_mean)
}
sim_results <- replicate(100000, simple_diff())
```

Resultados de Simulación

```r
mean(sim_results)
```

```
## [1] 0.605222
```



count
6000
4000
2000
0
−2
0
2
4
sim_results