Juan Menduiña

Trabajo Práctico N° 1: Modelo de Resultados Potenciales.

Ejercicio 1.

Considerar el ejemplo hipotético simple del Cuadro 1. Este ejemplo involucra una población de once pacientes, cada uno de los cuales está infectado con COVID-19. Hay dos tratamientos: ventiladores y reposo en cama. El Cuadro 1 muestra los resultados potenciales de cada paciente en términos de años de supervivencia después del tratamiento con cada tratamiento. Los valores de resultado más grandes corresponden a mejores resultados de salud.

Cuadro 1. *Resultados potenciales.*

Paciente	Y_T ***(u)***	Y_C ***(u)***	***TE***	***S***	***Y***
1	*1*	*10*	*-9*	*0*	*10*
2	*1*	*5*	*-4*	*0*	*5*
3	*1*	*4*	*-3*	*0*	*4*
4	*5*	*6*	*-1*	*0*	*6*
5	*5*	*1*	*4*	*1*	*5*
6	*9*	*7*	*2*	*1*	*9*
7	*6*	*8*	*-2*	*0*	*8*
8	*7*	*10*	*-3*	*0*	*10*
9	*8*	*2*	*6*	*1*	*8*
10	*9*	*6*	*3*	*1*	*9*
11	*10*	*7*	*3*	*1*	*10*

(a) *Calcular el efecto del tratamiento para cada paciente (columna TE).*

Cuadro 1.

(b) *¿Cuál es el efecto tratamiento promedio (ATE) para ventiladores (Y_T) comparado con reposo en cama (Y_C)? ¿Qué tipo de intervención es más efectiva en promedio?*

ATE= $\frac{\sum_{i=1}^{11} TE_i}{11}$

ATE= $\frac{-4}{11}$

ATE= -0,$\widehat{36}$.

Por lo tanto, el efecto tratamineto promedio (ATE) para ventiladores (Y_T) comparado con reposo en cama (Y_C) es -0,$\widehat{36}$. En promedio, la intervención reposo en cama es más efectiva.

(c) *Se supone que el "médico perfecto" conoce los resultados potenciales de cada paciente y, como resultado, elige el mejor tratamiento para cada paciente. Si se asigna a cada paciente el tratamiento más beneficioso para ese paciente, ¿qué pacientes recibirán* *ventiladores y cuáles recibirán reposo en cama (columna S)? Completar la última columna en función de lo que elija el médico perfecto.*

Cuadro 1.

(d) *¿Cuál es el efecto tratamiento promedio para ventiladores comparado con reposo en cama en el caso del médico perfecto (calculado con los datos observados)?*

ATE= $\frac{\sum_{i=1}^{11} Y_i}{11}$
ATE= $\frac{84}{11}$
ATE= $7,\widehat{63}$.

Por lo tanto, el efecto tratamiento promedio para ventiladores comparado con reposo en cama en el caso del médico perfecto (calculado con los datos observados) es $7,\widehat{63}$.

(e) *¿Cuál es la diferencia entre el ATE del médico perfecto y el ATE calculado en el inciso (b)? Explicar por qué el resultado es diferente en ambos casos y justificar cuál de los resultados será el correcto.*

La diferencia entre el ATE del médico perfecto y el ATE calculado en el inciso (b) es que el primero aparece un sesgo de selección cuando se calculan los resultados potencionales. El resultado correcto es el del inciso (b), ya que es el verdadero ATE.

Ejercicio 2.

En este ejercicio, se busca mostrar, utilizando simulaciones de Monte Carlo, que bajo el supuesto de independencia, la diferencia de medias para los individuos que recibieron el tratamiento y aquellos que no lo recibieron identifica el ATE. Considerar los siguientes datos.

Cuadro 2. *Resultados potenciales.*

Paciente	Y_T ***(u)***	Y_C ***(u)***
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*

(a) *Calcular el ATE.*

ATE= $\frac{\sum_{i=1}^{10} TE_i}{10}$
ATE= $\frac{6}{10}$
ATE= 0,6.

(b) *Generar una variable que, para cada observación, obtenga una realización de una normal estándar. Ordenar las observaciones de menor a mayor en función del valor de esta normal estándar.*

Stata.

(c) *Generar una variable d de otorgamiento del tratamiento que valga 1 para las primeras 5 observaciones ordenadas y que valga 0 para las restantes.*

Stata.

(d) *Computar la diferencia de medias en los promedios muestrales.*

Stata.

(e) *Repetir el procedimiento anterior 10000 veces y reportar la media de las diferencias de medias de cada simulación.*

Stata.

Ejercicio 3.

Sean Y_T, Y_C los resultados potenciales y D la variable de otorgamiento del tratamiento. Se define el efecto promedio de tratamiento como

ATE= E [Y_T - Y_C], **(1)**

el efecto promedio de tratamiento sobre los tratados como

ATT= E [Y_T - Y_C | D= 1], **(2)**

y el efecto promedio de tratamiento sobre los no tratados como

ATU= E [Y_T - Y_C | D= 0], **(3)**

En este ejercicio, se propone una descomposición del ATE diferente que la analizada en clase. La misma sigue el capítulo 4 de Causal Inference: The Mixtape de Scott Cunnningham.

(a) *Mostrar que ATE= πATT + (1 - π) ATU y describir los ponderadores π.*

ATE= E (Y_T - Y_C)
ATE= E (E (Y_T - Y_C) | D)
ATE= π E (Y_T - Y_C | D= 1) + (1 - π) E (Y_T - Y_C | D= 0)
ATE= πATT + (1 - π) ATU.

(b) *En base al libro mecionado arriba, interpretar que $\bar{y}_1$ - $\bar{y}_0$= ATE + Sesgo de Selección + Sesgo por heterogeneidad en tratamiento, en términos de efectos de tratamiento.*

$$\frac{1}{N_T}\sum_{i=1}^{n}(y_i|d_i = 1) - \frac{1}{N_C}\sum_{i=1}^{n}(y_i|d_i = 0) = E(Y^1) - E(Y^0) + E(Y^0 \mid D{=}1) - E(Y^0 \mid D{=}0) + (1 + \pi)(ATT - ATU).$$

- Sesgo de selección: Es una descripción de las diferencias entre los dos grupos si nunca se hubiera realizado un tratamiento.
- Sesgo por heterogeneidad en tratamiento: Es la diferencia en los retornos del tratamiento para los dos grupos multiplicada por la proporción de la población que está en el grupo de control.

(c) *Inicializar una muestra con 100 observaciones. Generar resultados potenciales de no recibir el tratamiento como $Y_0 \sim \mathcal{N}$ (100, 30).*

Stata.

(d) *Generar, ahora, un efecto de tratamiento constante e igual a 20 para todos, es decir,* TE_i *= 20 para todo i= 1, … , n. Generar una variable aleatoria normal estándar. Generar una variable de tratamiento* D_i *igual a 1 para aquellas observaciones que poseen un valor positivo en la variable aleatoria normal. Generar la variable Y observada como Y= D*Y_1 *+ (1 - D)* Y_0*. Computar la diferencia de medias y el test t. Luego, calcular ATE, ATT y ATU.*

Stata.

(e) *Repetir el inciso anterior, pero utilizar que TE* $\sim \mathcal{N}$ *(20, 10).*

Stata.

(f) *Repetir el inciso anterior, pero generar una variable aleatoria normal estándar rand y generar W= 1{rand* $>$ *0}. Generar el tratamiento como TE* $\sim \mathcal{N}$ *(20, 10) si W = 1 y TE* $\sim \mathcal{N}$ *(10, 10) si W= 0. Utilizar W como la variable utilizada para asignar el tratamiento.*

Stata.

(g) *Repetir el inciso anterior, pero, ahora, luego de generar los efectos de tratamiento en función de W, asignar el tratamiento aleatoriamente como en el primer inciso.*

Stata.

(h) *Comentar las conclusiones obtenidas con respecto a la heterogeneidad del tratamiento.*