Universidad de los Andes
Facultad de Economía

Capítulo 2

Definición de parámetros de impacto del tratamiento

Objetivo y problema

- Objetivo:

  Medir el impacto del tratamiento (p. ej., un programa) sobre una(s) variable(s) de resultado en un conjunto de individuos.

- Problema:

  Para obtener el efecto del tratamiento, necesitaríamos evaluar la diferencia entre la variable de resultado de los participantes con programa y la variable de resultado de los participantes en ausencia del programa, denominado *contrafactual*.

Notación

$$D_i = \begin{cases} 1 & \text{individuo } i \text{ recibe tratamiento} \\ 0 & \text{de lo contrario} \end{cases}$$

Definiciones: Grupo de tratamiento: Grupo de individuos participantes en el tratamiento. Grupo de control: Grupo de individuos no participantes en el tratamiento.

$Y_i(D_i)$ { variable de resultado para cada individuo i dado su estado D_i

$Y_i(1)$ es la variable de resultado del individuo i si es tratado
$Y_i(0)$ es la variable de resultado del individuo i si no es tratado.

Ejemplo *Canasta*

Suponga que debemos evaluar el impacto del programa *Canasta*, que es un programa de nutrición.

- Población objetivo: Niños entre 0 y 6 años pertenecientes a Sisben1 y 2.
- Consiste en: Proveer un mercado mensual por el valor de $X, entregado a las madres de los niños participantes.

Ejemplo *Canasta*

Indicador de participación definido como:

$$D_i = \begin{cases} 1 \text{ si el niño elegible participa en el programa de nutrición} \\ 0 \text{ si el niño elegible no participa en el programa de nutrición} \end{cases}$$

Variables de resultado:

$$Y_i = \{ \text{Talla según la edad}$$

Efecto del tratamiento

El efecto del tratamiento o programa sobre el individuo i estaría dado por:

$$\boxed{\tau_i = Y_i(1) - Y_i(0)}$$

Problema: Sólo se puede observar **uno** de los dos resultados potenciales para cada individuo i . El individuo i sólo puede ser participante o no participante, pero no ambas al tiempo.

El resultado efectivamente observado se puede escribir:

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

Dado que no se puede obtener el efecto del tratamiento para cada individuo i. El análisis debe concentrarse en el impacto promedio del programa en la población o en subconjuntos de la población.

¿Qué se puede estimar?: ATE

ATE: Efecto promedio del programa

$$\boxed{\tau_{ATE} = E(\tau_i) = E[Y_i(1) - Y_i(0)]}$$

- Se interpreta como el cambio en la variable resultado cuando un individuo pasa aleatoriamente de ser participante a ser no participante.
- Relevante particularmente para la evaluación de programas universales.

¿Qué se puede estimar?: ATT

ATT: Efecto promedio del programa sobre los tratados

Defina:

$E[Y_i(1)|D_i = 1]$ como el promedio de la variable de resultado de individuos tratados dado que participaron en el programa.

$E[Y_i(0)|D_i = 1]$ como el promedio de la variable de resultado de individuos tratados si no hubieran participado en el tratamiento. Es decir, cómo les habría ido en el escenario hipotético de que el programa no hubiera existido.

$\Downarrow$

Resultado CONTRAFACTUAL / no se observa

¿Qué se puede estimar?: ATT

ATT: Efecto promedio del programa sobre los tratados

$$\tau_{ATT} = E(\tau|D=1) = E[Y(1)|D=1] - E[Y(0)|D=1]$$

CONTRAFACTUAL

Mide el impacto del programa sobre los tratados (el subconjunto de individuos que fueron efectivamente tratados).

Parámetro de mayor interés en la evaluación de impacto, porque la mayoría de programas son focalizados y no universales.

¿Qué se puede estimar?: ATU

ATU: Efecto promedio del programa sobre los no tratados

Defina:

$E[Y_i(0)|D_i = 0]$ como el promedio de la variable de resultado de individuos no tratados dado que no participaron en el programa.

$E[Y_i(1)|D_i = 0]$ como el promedio de la variable de resultado de individuos no tratados si hubieran podido participar en el tratamiento.

Resultado CONTRAFACTUAL / no se observa

¿Qué se puede estimar?: ATU

ATU: Efecto promedio del tratamiento sobre los no tratados

$$\tau_{ATU} = E(\tau|D=0) = E[Y(1)|D=0] - E[Y(0)|D=0]$$

CONTRAFACTUAL

Parámetro relevante para determinar si el programa se debe extender a otros subconjuntos de población fuera del que en la actualidad es elegible.

Aproximación al *contrafactual*

Para calcular ATT y ATU se debe encontrar una aproximación para el *contrafactual*.

Para ATT:

No observado — Observado

$$E[Y(0)|D=1] \overset{?}{\approx} E[Y(0)|D=0]$$

→ Podría ser una aproximación inexacta, dado que quienes participan en el programa, en general, son distintos de quienes no lo hacen; por lo tanto, la variable de resultado suele ser distinta entre el grupo de control y el de tratamiento, aun en ausencia del programa.

Condición para la aproximación

Para recuperar τ_{ATT} se requiere que:

$$E[Y(0)|D=1]-E[Y(0)|D=0]$$

Implica que la variable de resultado en ausencia del programa debe ser idéntica para los individuos tratados y los individuos no tratados.
En otras palabras, la participación en el programa es independiente de las características de los individuos.

Universidad de los Andes
Facultad de Economía

ATT aproximado

Asumiendo que se cumple:

$$E[Y(0)|D=1]-E[Y(0)|D=0]=0$$

τ_{ATT} se puede expresar como:

$$\tau_{ATT}=E[Y(1)|D=1]-E[Y(0)|D=0]$$

Análogo muestral de ese momento poblacional es:

$$\tau_{ATT}=[\bar{Y}|D=1]-[\bar{Y}|D=0]$$

Donde $\bar{Y}$ es el promedio muestral de la variable de resultado dado su estatus de participación.

Universidad de los Andes
Facultad de Economía

Independencia condicional

El estimador τ_{ATT} se puede obtener de la siguiente regresión:

$$Y_i = \beta_0 + \beta_1 D_i + u_i$$

Si la condición:

$$E[Y(0)|D = 1] - E[Y(0)|D = 0] = 0$$

se cumple, esto implica que:

$$\underbrace{E[u_i|D_i] = 0}_{\text{Independencia condicional.}}$$

Quienes participan en el programa no son sistemáticamente distintos de aquellos que no participan en el programa.

Universidad de los Andes
Facultad de Economía

Propiedades del estimador

Bajo independencia condicional el estimador $\hat{\beta}_1$ de mínimos cuadrados ordinarios (MCO) de es:

1. Consistente:

$$\widehat{\beta_1} \xrightarrow{p} \beta_1$$

2. Insesgado:

$$E(\widehat{\beta_1}) = \beta_1$$

Universidad de los Andes
Facultad de Economía

¿Qué es β_1?

β_1
- Efecto del programa
- Diferencia de medias de la variable de resultado entre el grupo de tratamiento y el grupo de control

$$\tau_{ATT} = E[Y(1)|D=1] - E[Y(0)|D=0]$$
$$= \beta_0 + \beta_1 - \beta_0 = \beta_1$$

Por lo tanto, el estimador de β_1 estaría dado por:

$$\widehat{\beta}_1 = [\bar{Y}|D=1] - [\bar{Y}|D=0]$$

Ejemplo *Canasta*

$E[Y(1)|D = 1]$ → Corresponde al promedio de la talla según la edad entre los participantes, en presencia del programa.

$E[Y(0)|D = 1]$ → Corresponde al promedio de la talla según la edad entre los participantes en ausencia del programa. El contrafactual.

$E[Y(0)|D = 0]$ → Corresponde al promedio de la talla según la edad en el grupo de no participantes. Promedio en el grupo de control.

Ejemplo *Canasta*

- Si se cumple que:

$$E[Y(0)|D=1]-E[Y(0)|D=0]=0$$

Entonces el efecto del programa se podría estimar a través de la comparación de medias de la variable de resultado del grupo de tratamiento y el grupo de control.

El efecto se puede estimar con la regresión

$$Y_i=\beta_0+\beta_1 D_i+u_i$$

Efecto del programa

Capítulo 3

Sesgo de selección

Notación

$$D_i = \begin{cases} 1 & \text{individuo } i \text{ recibe tratamiento} \\ 0 & \text{de lo contrario} \end{cases}$$

$Y_i(D_i)$ { Variable de resultado para cada individuo i dado su estado D_i

$Y_i(1)$ es la variable de resultado del individuo i si es tratado

$Y_i(0)$ es la variable de resultado del individuo i si no es tratado

Efecto individual del programa: $\tau_i = Y_i(1) - Y_i(0)$

Notación

$E[Y(1)|D=1] \longrightarrow$ Media de la variable de resultado en el grupo de participantes.

$E[Y(0)|D=1] \longrightarrow$ Media que hubieran obtenido los participantes si el programa no hubiera existido.

$E[Y(1)|D=0] \longrightarrow$ Media de la variable de resultado entre los no participantes en el programa, si hubieran podido participar.

$E[Y(0)|D=0] \longrightarrow$ Media de la variable de resultado de quienes no participaron en el programa.

Reto de la evaluación

El objeto de la evaluación de impacto es estimar:

$$\tau_{ATT} = E[\tau|D=1] = E[Y(1)|D=1] - \underbrace{E[Y(0)|D=1]}_{\text{No observado}}$$

Para calcularlo podemos suponer que:

$$E[Y(0)|D=1] - E[Y(0)|D=0] = 0$$

El principal reto es determinar bajo qué condiciones este supuesto sería válido.

El supuesto se viola cuando $\longrightarrow$ hay autoselección.

Sesgo de selección

Cuando los individuos se autoseleccionan en el programa:

– Existen características observadas y no observadas que causan que unos individuos decidan participar y otros decidan no hacerlo. Por lo tanto, los dos grupos podrían ser sistemáticamente diferentes.

  (Características *no observadas* son aquellas que no se miden o no quedaron registradas en la base de datos)

– El grupo de tratamiento y el de control son generalmente distintos, aun en ausencia del programa.

Sesgo de selección

$$E[Y(0)|D=1] \neq E[Y(0)|D=0]$$

Ejemplo *Canasta*

Suponga que en el programa *Canasta* los individuos deben decidir si participar o no (p. ej., el programa no es obligatorio).

Participar tiene un costo:

- Tiempo
- Trámites

P. ej., formularios, controles médicos, etc.

Lo que puede generar que existan diferencias entre el perfil de las madres de los niños elegibles, que aplican y las que no aplican para participar.

P. ej.: Madres más o menos proactivas.

Puede incidir en el estado nutricional de los niños.

Consecuencia del sesgo de selección

- Al violarse el supuesto:

$$E[Y(0)|D=1]-E[Y(0)|D=0] \neq 0$$

$$\Downarrow$$

en la regresión $Y_i = \beta_0 + \beta_1 D_i + u_i$ implica:

$$E[u_i|\, D_i] \neq 0$$

Existen características observadas y no observadas de los individuos contenidas en u_i que explican tanto la decisión de participar en el programa, D_i , como la variable de resultado Y_i.

Consecuencias en las propiedades del estimador

- Cuando se viola el supuesto de independencia condicional, el estimador $\hat{\beta}_1$ de mínimos cuadrados ordinarios (MCO) deja de ser insesgado.

⇩

- Por ende, el efecto del programa no se puede obtener de la comparación de las medias de la variable de resultado entre los tratados y no tratados. $\hat{\beta}_1$ captura tanto el efecto directo del programa como el efecto que tienen otros factores no medidos (p. ej., proactividad y motivación de la madre) sobre el estado nutricional del niño… y ¡no se pueden separar!

¿Cómo eliminar el sesgo? Variables observadas

- Si las características del individuo que explican tanto la participación en el programa como la relación con la variable de resultado son observadas y están contenidas en la base de datos, se deben introducir explícitamente en la regresión, para eliminar el sesgo, así:

$$Y_i = \beta_0 + \beta_1 D_i + \underbrace{\beta_2 X_i}_{\text{Características observadas del individuo } i.} + u_i$$

Ejemplo *Canasta*

Si los individuos más pobres son aquellos que deciden participar en el programa, mientras que los más ricos deciden no hacerlo

Entonces la variable de control incluida en la regresión:

$X_i \longrightarrow$ es el ingreso de las familias

De esta manera se elimina el sesgo de selección y $\hat{\beta}_1$ es un estimador insesgado y consistente del programa.

Variables no observadas

- Si las variables que determinan la participación en el programa no se observan o no están registradas en la base $\Rightarrow$

$$E(\widehat{\beta_1}) \neq \beta_1$$

Suponiendo que w_i es la variable omitida, es decir:

$$Y_i = \beta_0 + \beta_1 D_i + \underbrace{\beta_2 W_i}_{\text{Variable no observada}} + u_i$$

Al no poder incluir w en la regresión, le estamos atribuyendo al programa parte del efecto que tiene w sobre Y.

Ejemplo *Canasta*

Si, por ejemplo, la diferencia entre los niños participantes y no participantes es la dedicación de la madre

y

La dedicación de la madre no es una variable medida y registrada dentro de su encuesta, entonces:

⇩

$$E(\widehat{\beta_1}) \neq \beta_1$$

Dirección del sesgo

- La dirección del sesgo depende de la correlación entre la variable no observada (omitida) y la participación del individuo en el programa:

	$Corr(\mathrm{D_i W_i}) > 0$	$Corr(\mathrm{D_i W_i}) < 0$
$\beta_2 > 0$	$E(\widehat{\beta_1}) > \beta_1$	$E(\widehat{\beta_1}) < \beta_1$
$\beta_2 < 0$	$E(\widehat{\beta_1}) < \beta_1$	$E(\widehat{\beta_1}) > \beta_1$

Ejemplo *Canasta*

Si la dedicación de la madre

aumenta la probabilidad de participar en el programa → $Corr(\mathrm{D_i W_i}) > 0$

y aumenta el peso por talla del niño → $\beta_2 > 0$

$$E(\widehat{\beta_1}) > \beta_1$$

Si la dedicación de la madre

disminuye la probabilidad de participar en el programa → $Corr(\mathrm{D_i W_i}) < 0$

pero aumenta el peso por talla del niño → $\beta_2 > 0$

$$E(\widehat{\beta_1}) < \beta_1$$