

Clase 22: Resultados Potenciales en Regresiones y Ejemplos

Haciendo Economía I
Econ 2205

Ignacio Sarmiento-Barbieri

Universidad de los Andes

October 29, 2023

- ▶ Las referencias para esta clase siguen siendo:
 - 1 Mastering Metrics de Angrist y Pischke (cap 1)
 - 2 Mostly Harmless Econometrics de Angrist y Pischke (cap 2)
 - 3 Causal Inference: The Mixtape de Cunningham (cap 4) (disponible online en su pagina web)
- ▶ Quiz sobre Resultados Potenciales → Martes 7 de Noviembre

Plan para hoy

- 1 Anuncios
- 2 Modelo de Resultados Potenciales Recap
 - Ejemplo: El Experimento STAR
- 3 Análisis de Regresión en Experimentos
 - Ejemplo: El Experimento STAR (cont.)

Modelo de Resultados Potenciales

- ▶ Una variable binaria de tratamiento (e.g., va al hospital, no va):

$$D_i = \{0, 1\} \quad (1)$$

- ▶ Resultado (Outcome) para el individuo i (e.g., salud):

$$Y_i \quad (2)$$

- ▶ El modelo tiene varios nombres...
 - ▶ Splawa-Neyman marco de resultados potenciales
 - ▶ Rubin modelo causal
 - ▶ Neyman-Rubin (Splawa-Neyman-Rubin)

Modelo de Resultados Potenciales

- ▶ Pregunta de investigación: D_i afecta a Y_i ?
- ▶ Para cada individuo i , hay dos **resultados potenciales** (con D_i binario)
 - ▶ Y_{1i} si $D_i = 1$ Resultado de i si va al hospital
 - ▶ Y_{0i} si $D_i = 0$ Resultado de i si no va al hospital
- ▶ La diferencia entre los dos resultados nos da **el efecto causal del tratamiento de ir al hospital**, i.e.,

$$\tau_i = Y_{1i} - Y_{0i} \quad (3)$$

Modelo de Resultados potenciales

Problemas

- ▶ Esta ecuación simple:

$$\tau_i = Y_{1i} - Y_{0i} \quad (4)$$

- ▶ nos lleva al **problema fundamental de inferencia causal**.
- ▶ Nunca podemos observar simultáneamente Y_{1i} and Y_{0i} .



Source: <https://tinyurl.com/yv3v5tum>

Modelo de Resultados potenciales

Solución Propuesta

- ▶ Comparamos
 - ▶ resultado para los que van al hospital ($Y_{1i} \mid D_i = 1$)
 - ▶ resultados para los que **no** van al hospital ($Y_{0i} \mid D_i = 0$)

$$E[Y_i \mid D_i = 1] - E[Y_i \mid D_i = 0] \quad (5)$$

- ▶ que nos da la *diferencia observada en los resultados de salud*

Modelo de Resultados potenciales

Solución Propuesta

- ▶ Entonces

$$E[Y_i | D_i = 1] - E[Y_i | D_i = 0] = E[Y_{1i} | D_i = 1] - E[Y_{0i} | D_i = 0] \quad (6)$$

- ▶ Haciendo un poco de matemática creativa

$$= E[Y_{1i} | D_i = 1] - E[Y_{0i} | D_i = 1] + E[Y_{0i} | D_i = 1] - E[Y_{0i} | D_i = 0] \quad (7)$$

Modelo de Resultados potenciales

Solución Propuesta

Que nos dice esto entonces?

$$E[Y_i | D_i = 1] - E[Y_i | D_i = 0] = \quad (8)$$

$$\underbrace{E[Y_{1i} | D_i = 1] - E[Y_{0i} | D_i = 1]}_{\text{Efecto Promedio sobre los tratados (ATT) ☺}} + E[Y_{0i} | D_i = 1] - E[Y_{0i} | D_i = 0] \quad (9)$$

Modelo de Resultados potenciales

Solución Propuesta

Que nos dice esto entonces?

$$E[Y_i | D_i = 1] - E[Y_i | D_i = 0] = \quad (10)$$

$$= \underbrace{E[Y_{1i} | D_i = 1] - E[Y_{0i} | D_i = 1]}_{\text{Efecto Promedio sobre los tratados (ATT) ☺}} + \underbrace{E[Y_{0i} | D_i = 1] - E[Y_{0i} | D_i = 0]}_{\text{Sesgo de Selección ☺}} \quad (11)$$

Modelo de Resultados potenciales

Solución Propuesta

- ▶ El **primer término** es la *variación buena* que es la respuesta que queremos

$$E[Y_{1i} \mid D_i = 1] - E[Y_{0i} \mid D_i = 1] \quad (12)$$

$$= E[Y_{1i} - Y_{0i} \mid D_i = 1] \quad (13)$$

$$= E[\tau_i \mid D_i = 1] \quad (14)$$

- ▶ El **efecto causal promedio** de la hospitalización para los *individuos hospitalizados* (ATT).

Modelo de Resultados potenciales

Solución Propuesta

- ▶ El **segundo término** es la *variación mala* que no nos dejan obtener la respuesta que queremos

$$E[Y_{0i} \mid D_i = 1] - E[Y_{0i} \mid D_i = 0] \quad (15)$$

- ▶ La diferencia en el promedio de resultados para los no tratados entre el tratamiento y el control.
- ▶ **Sesgo de selección** cuan malo es el grupo de control como contra factual.

Modelo de Resultados potenciales

Ejemplo: Programas de entrenamiento

- ▶ Los gobiernos suelen hacer programas de entrenamiento para ayudar a trabajadores necesitados
- ▶ La pregunta es: Estos programas tienen los efectos deseados (por ej. subir salarios)?
- ▶ Los estudios observacionales que comparan los datos de salarios de participantes y no participantes suelen encontrar que aquellos que terminan estos programas terminan con salarios mas bajos.
- ▶ El problema es que los participantes se auto seleccionan
- ▶ y que los participantes suelen tener salarios mas bajos

Modelo de Resultados potenciales

Ejemplo: Programas de entrenamiento

Nearly All Programs Track Multiple Outcome Measures, but Little is Known about Program Effectiveness

Almost all programs tracked multiple outcome measures related to employment and training, and many programs tracked similar measures. Forty-one of the 47 programs tracked at least three outcome measures in fiscal year 2009, according to officials. The most frequently tracked outcome measure was “entered employment”—the number of program participants who found jobs (see table 2). Many programs also tracked “employment retention” and “wage gain or change.” These are the types of measures developed under the Office of Management and Budget’s (OMB) common measures initiative, which sought to unify definitions for performance across programs with similar goals.¹⁹ Three programs did not track any outcome measures at the federal level in fiscal year 2009.²⁰ For a detailed list of outcome measures tracked by federal employment and training programs, see appendix V.

Table 2: Outcome Measures Tracked Most Frequently by Programs in Fiscal Year 2009

Outcome measures	Number of programs measuring this outcome
Entered employment	38
Employment retention	29
Wage gain or change	23
Credential attainment	19
Other “positive outcomes” ^a	17
Educational attainment	16
Customer satisfaction	8
Other outcomes ^b	23
No outcome measures	3

Source: Government Accountability Office, “Multiple Employment and Training Programs,” GAO-11-92, January 2011, p. 11.

Modelo de Resultados potenciales

Ejemplo: Programas de entrenamiento

Downsizing the Federal Government

YOUR GUIDE TO CUTTING FEDERAL SPENDING

HOME

ABOUT

BLOG

FEDERAL POLICY BASICS

SPENDING CHARTS

FEDERAL DEPARTMENTS

Agriculture

Commerce

Defense

Employment and Training Programs: Ineffective and Unneeded

by Chris Edwards and Daniel J. Murphy

June 1, 2011

Source: <https://www.downsizinggovernment.org/labor/employment-training-programs>

Modelo de Resultados potenciales

Ejemplo: Programas de entrenamiento

- ▶ Los gobiernos suelen hacer programas de entrenamiento para ayudar a trabajadores necesitados
- ▶ La pregunta es: Estos programas tienen los efectos deseados (por ej. subir salarios)?
- ▶ Los estudios observacionales que comparan los datos de salarios de participantes y no participantes suelen encontrar que aquellos que terminan estos programas terminan con salarios mas bajos.
- ▶ El problema es que los participantes se auto seleccionan
- ▶ y que los participantes suelen tener salarios mas bajos

Modelo de Resultados potenciales

Ejemplo: Programas de entrenamiento

- ▶ Como formalizamos esta preocupación en nuestro modelo?
- ▶ Evaluación observacional de programas

$$E[\text{Salario}_i \mid \text{Programa}_i = 1] - E[\text{Salario}_i \mid \text{Programa}_i = 0] = \quad (16)$$

$$\underbrace{E[\text{Salario}_{1i} \mid \text{Programa}_i = 1] - E[\text{Salario}_{0i} \mid \text{Programa}_i = 1]}_{\text{Efecto causal promedio del programa de entrenamiento en los salarios de los participantes i.e., } \bar{\tau}} + \quad (17)$$

Efecto causal promedio del programa de entrenamiento en los salarios de los participantes i.e., $\bar{\tau}$

$$\underbrace{E[\text{Salario}_{0i} \mid \text{Programa}_i = 1] - E[\text{Salario}_{0i} \mid \text{Programa}_i = 0]}_{\text{Sesgo de Selección}} \quad (18)$$

Modelo de Resultados Potenciales

Ejemplo: Tamaño de la clase y aprendizaje de estudiantes

- ▶ Uno de los insumos más costoso en la función de producción educativa es el tamaño de la clase:
- ▶ Clases mas pequeñas → mas maestros
- ▶ Fundamental entender el beneficio de tamaño de la clase más pequeño
- ▶ Estudios **observacionales** (no experimentales) sugieren que existe poca o ninguna relación entre el tamaño de la clase y el aprendizaje de los estudiantes.
- ▶ De ser cierto, esto implicaría que se puede ahorrar dinero contratando menos maestros sin la consecuente reducción en el rendimiento.

Modelo de Resultados Potenciales

Ejemplo: Tamaño de la clase y aprendizaje de estudiantes

- ▶ Como formalizamos esta preocupación en nuestro modelo?
- ▶ Evaluación observacional de programas

$$E \left[\text{Calificacion}_i \mid \text{Clase Pequeña}_i = 1 \right] - E \left[\text{Calificacion}_i \mid \text{Clase Pequeña}_i = 0 \right] = \quad (19)$$

Modelo de Resultados Potenciales

Experimentos

- ▶ Como hacen los experimentos para resolver el sesgo de selección?

Modelo de Resultados Potenciales

Experimentos

- ▶ Como hacen los experimentos para resolver el sesgo de selección?
- ▶ Los experimentos rompen el link entre los resultados potenciales y el tratamiento

Modelo de Resultados Potenciales

Experimentos

- ▶ Como hacen los experimentos para resolver el sesgo de selección?
- ▶ Los experimentos rompen el link entre los resultados potenciales y el tratamiento
- ▶ En otras palabras, asignando aleatoriamente D_i hace que D_i sea independiente de que resultado observamos (es decir Y_{1i} o Y_{0i})

$$\{Y_{1i}, Y_{0i}\} \perp D_i \quad (20)$$

- ▶ La asignación aleatoria implica que los grupos de tratamiento y control salen de la misma población subyacente.
- ▶ Son iguales en todos los sentidos, incluyendo sus $E[Y_{0i}]$

Modelo de Resultados Potenciales

Experimentos

Ejemplo: Tamaño de la clase y aprendizaje de estudiantes

El Experimento STAR

- ▶ Estudios **observacionales** (no experimentales) sugieren que existe poca o ninguna relación entre el tamaño de la clase y el aprendizaje de los estudiantes.
- ▶ De ser cierto, esto implicaría que se puede ahorrar dinero contratando menos maestros sin la consecuente reducción en el rendimiento.
- ▶ Problema: estudiantes más débiles suelen ser colocados en clases más pequeñas

Ejemplo: Tamaño de la clase y aprendizaje de estudiantes

El Experimento STAR

- ▶ Estudios **observacionales** (no experimentales) sugieren que existe poca o ninguna relación entre el tamaño de la clase y el aprendizaje de los estudiantes.
- ▶ De ser cierto, esto implicaría que se puede ahorrar dinero contratando menos maestros sin la consecuente reducción en el rendimiento.
- ▶ Problema: estudiantes más débiles suelen ser colocados en clases más pequeñas
- ▶ Un **experimento** aleatorio puede solucionar esto: aleatorizar estudiantes a clases de diferentes tamaños
- ▶ Esta es la idea del proyecto STAR de Tennessee

Ejemplo: Tamaño de la clase y aprendizaje de estudiantes

El Experimento STAR

- ▶ El experimento STAR fue ambicioso e influyente
- ▶ El experimento asignó a los estudiantes a uno de tres tratamientos:
 - 1 Clases **pequeñas** con 13 – 17 niños,
 - 2 Clases **regulares** con 22 – 25 estudiantes y un profesor asistente tiempo parcial (**grupo control**),
 - 3 Clases **regulares** con un profesor asistente de tiempo completo
- ▶ Costó alrededor de \$12 millones y se implementó para una cohorte de niños de jardín de infantes en 1985/86.
- ▶ El estudio duró cuatro años e involucró a unos 11.600 niños.

Ejemplo: Tamaño de la clase y aprendizaje de estudiantes

El Experimento STAR

- Lo *primero* que tenemos que preguntarnos: Funcionó la aleatorización?

Variable	Pequeña	Regular	Regular con Asistente	P-value conjunto
Almuerzo Gratis	0.47	0.48	0.50	0.09
Blanco/Asiático	0.68	0.67	0.66	0.26
Edad en 1985	5.44	5.43	5.42	0.32

Ejemplo: Tamaño de la clase y aprendizaje de estudiantes

El Experimento STAR

- Lo *primero* que tenemos que preguntarnos: Funcionó la aleatorización?

Variable	Pequeña	Regular	Regular con Asistente	P-value conjunto
Almuerzo Gratis	0.47	0.48	0.50	0.09
Blanco/Asiático	0.68	0.67	0.66	0.26
Edad en 1985	5.44	5.43	5.42	0.32
Tamaño de clase	15.10	22.40	22.80	0.00

Ejemplo: Tamaño de la clase y aprendizaje de estudiantes

El Experimento STAR

- Lo *primero* que tenemos que preguntarnos: Funcionó la aleatorización?

Variable	Pequeña	Regular	Regular con Asistente	P-value conjunto
Almuerzo Gratis	0.47	0.48	0.50	0.09
Blanco/ Asiático	0.68	0.67	0.66	0.26
Edad en 1985	5.44	5.43	5.42	0.32
Tamaño de clase	15.10	22.40	22.80	0.00
Percentil Resultado de Prueba	54.70	48.90	50.00	0.00

Fuente: Adaptación tabla 2.2.1 MHE

1 Anuncios

2 Modelo de Resultados Potenciales Recap

- Ejemplo: El Experimento STAR

3 Análisis de Regresión en Experimentos

- Ejemplo: El Experimento STAR (cont.)

Análisis de Regresión en Experimentos

- ▶ La tabla previa estima/compara el efecto del tratamiento (ATE) haciendo diferencias de medias .
- ▶ Podemos hacer lo mismo con una regresión.
- ▶ Específicamente, si hacemos la regresión del "outcome" (percentil de la prueba) en una variable dummy para cada grupo de tratamiento.

Análisis de Regresión en Experimentos

- ▶ Asumimos que el efecto del tratamiento es constante ("homogeneo")