

Ejercicio Ejemplo 1: Programa de Capacitación Laboral

Se ofreció un programa de capacitación laboral a 500 personas desempleadas, de las cuales 300 eligieron participar. Seis meses después, 210 de los que participaron habían encontrado empleo, mientras que solo 100 de los que no participaron consiguieron trabajo.

Preguntas:

1. Define los resultados potenciales Y^1 y Y^0 en este contexto, en lenguaje sencillo, refiriéndote a los resultados laborales.
2. Calcula el efecto del tratamiento ingenuo.
3. ¿Por qué podría estar sesgado este cálculo?
4. ¿Qué supuestos deberían cumplirse para que esto represente el efecto causal?
5. Si la motivación para encontrar trabajo afecta tanto la participación en el programa como los resultados laborales, explica cómo esto crea un sesgo de selección en términos de resultados potenciales.
6. ¿Qué esperarías que sea mayor en este contexto: el ATE o el ATT? Justifica basándote en razonamiento económico.

Ejercicio 2: Cálculo de Efectos del Tratamiento con Selección Observable

Considera una población de 6 individuos con los siguientes resultados potenciales para ingresos (en miles de dólares) bajo tratamiento (programa de capacitación laboral) y control (sin capacitación):

Persona	Y^0 - Sin Capacitación	Y^1 - Con Capacitación
1	20	21
2	22	24
3	18	21
4	30	34
5	25	30
6	15	21

Supón que antes del tratamiento, todos los individuos tienen ingresos Y^0 , es decir, no hay efecto por no recibir el tratamiento, y Y^0 es conocido.

Preguntas:

1. Supón que hay recursos para capacitar a 3 individuos, y se ofrece capacitación a los individuos con menor ingreso actual. Identifica qué individuos recibirían el tratamiento bajo esta regla. Luego calcula:
 - a) El Efecto Promedio del Tratamiento (ATE) para toda la población
 - b) El Efecto Promedio del Tratamiento sobre los Tratados (ATT)
 - c) El Efecto Promedio del Tratamiento sobre los No Tratados (ATU)
2. Calcula ahora el estimador ingenuo de diferencia de medias (promedio del resultado para los tratados menos promedio del resultado para los no tratados).
3. Compara el estimador ingenuo con el ATE verdadero. ¿Es una sobreestimación o subestimación? Explica por qué ocurre esto en términos de sesgo de selección.
4. Para esta regla de selección y este conjunto de resultados potenciales, ¿el sesgo de selección tiende a hacer que el programa parezca más o menos efectivo de lo que realmente es? Proporciona una intuición económica de por qué esto podría ocurrir en programas de capacitación laboral reales.
5. Supón ahora que la regla de selección fuese ofrecer capacitación a los 3 individuos que más se benefician de ella. ¿Qué individuos recibirían el tratamiento?
 - ¿Cuál sería el ATT y el ATU bajo esta regla?
 - ¿Cómo se compararía el estimador ingenuo con el ATE ahora?
 - ¿Por qué no podemos implementar esta regla de selección en la práctica?

Ejercicio 3: Ensayo sobre colonoscopia

En 2022, el *New England Journal of Medicine* publicó [un estudio](#) sobre los beneficios para la salud de las colonoscopías. Las colonoscopías pueden ayudar a detectar etapas tempranas de cáncer, pero son muy incómodas y conllevan ciertos riesgos, por lo que solo deberían recomendarse si los beneficios son suficientemente grandes. El estudio utilizó un diseño de incentivo: de 84 585 participantes, todos entre 55 y 64 años, a 28 220 se les envió una invitación para realizarse el examen. En el grupo invitado, 11 843

pacientes se sometieron a la prueba y 0 en el grupo de control. Tras 10 años, la incidencia de cáncer colorrectal fue de 259 en el grupo invitado y 622 en el grupo de control.

1. Describa en este experimento cuál es la variable de asignación del tratamiento (Z), la variable de tratamiento (D) y el resultado (Y).
2. Calcule el ITT y expréselo en puntos porcentuales absolutos y en términos relativos. Escriba un párrafo muy corto para un periódico explicándolo en el contexto de este experimento (máximo 100 palabras).
3. Cuando se publicó este artículo, muchos científicos lo consideraron un argumento en contra de recomendar colonoscopías. Dado los riesgos y costos, sostuvieron que el efecto causal que calculó en la parte 2 era demasiado pequeño para justificar el procedimiento. Sin embargo, otros señalaron que el incentivo no indujo mucha participación, por lo que era esperable que el efecto fuese pequeño. Argumentan que, una vez que se ajusta por la baja participación, el efecto real es suficiente para recomendar la prueba. Calcule el efecto al que se refiere este segundo grupo y explique por qué es un número relevante para informar la política pública.
4. Algunos científicos respondieron que el número que calculó en el punto 3 está sesgado al alza, porque las personas que aceptan hacerse colonoscopías probablemente estén en mayor riesgo. Argumentan que, si se diera una colonoscopia a una persona aleatoria de 55 a 64 años, los efectos serían probablemente menores y, por lo tanto, este no es el efecto clínicamente relevante. Explique a qué se refieren, usando el marco de resultados potenciales.
5. ¿Con qué grupo está de acuerdo? Escriba un párrafo breve para un periódico defendiendo la relevancia del número que eligió (máximo 100 palabras).

Ejercicio 4: Programa de Tutoría de Verano: Evaluación del Efecto de Tratamiento

Una gran ciudad llevó a cabo un **programa de tutoría de verano** para estudiantes que ingresarían al 10° grado, con el objetivo de mejorar sus habilidades matemáticas antes de comenzar el nuevo año escolar.

- **Aleatorización.** 1,000 estudiantes voluntarios fueron asignados aleatoriamente:

- 500 al grupo de **Tratamiento** (se les ofreció el campamento gratuito de tutoría de cuatro semanas)
- 500 al grupo de **Control** (sin oferta)
- **Resultado de interés.** Una prueba estandarizada de matemáticas administrada **ocho meses después** (abril del siguiente año escolar).
- **Deserción selectiva.** En abril, **algunos estudiantes no realizaron la prueba de seguimiento**:
 - Algunos estudiantes simplemente no se presentaron el día del examen. Debido a que estos "ausentes" nunca realizaron el examen, su resultado está ausente en nuestro conjunto de datos.

Te han contratado para estimar el efecto del tratamiento. Los datos están disponibles en [este enlace](#).¹

Parte A: Estimación ingenua

1. Carga los datos y conserva solo las filas donde se observa el resultado.
2. Estima el efecto del tratamiento sin controles, y controlando por la puntaje previa.
3. Interpreta los resultados. ¿Qué suposición sobre la deserción estás haciendo *implícitamente* y por qué podría fallar en este caso?

Parte B: Diagnóstico del problema

1. Calcula las tasas de deserción en cada grupo. Repórtalas claramente.
2. Explica, con palabras, *cómo* la deserción diferencial puede sesgar la estimación ingenua.
3. Supongamos que la deserción fuera completamente aleatoria. ¿Sería insesgado el estimador en **A2**? Justifica brevemente. ¿Y si fuera aleatoria condicionando en la puntaje previa?

Parte C: Límites de Lee (2009) El método de Lee proporciona un **intervalo** para el efecto promedio del tratamiento.

1. Calcula las tasas de retención. ¿Cuál es la tasa de deserción diferencial? ¿Qué grupo tiene más deserción?
2. Queremos saber cuáles serían los resultados en el grupo con menos deserción si tuviera la misma tasa que el grupo con más deserción. Para hacer esto, eliminaremos individuos de la muestra para igualar los niveles. (Deberían tener el

¹ Puedes cargarlos directamente haciendo:

```
data <- read.csv("https://jmbvgarcia.github.io/assets/code/datos_ps2.csv")
```

mismo número de individuos después de esto). Haz esto de dos maneras diferentes:

1. Suponiendo que los estudiantes que abandonan son los de peor rendimiento. Es decir, elimina a los estudiantes con las puntuaciones más bajas. ¿Cuál sería el efecto estimado entonces?
2. Suponiendo que los estudiantes que abandonan son los de mejor rendimiento. Es decir, elimina a los estudiantes con las puntuaciones más altas. ¿Cuál sería el efecto estimado entonces?
3. Reporta los dos efectos estimados. Este método se llama límites de Lee y puede utilizarse en casos donde la deserción no es demasiado fuerte. Escribe un párrafo breve resumiendo tus hallazgos para el secretario de educación responsable del programa.
4. Basándote en el contexto, ¿esperas que el efecto verdadero esté más cerca del límite superior o del límite inferior? Justifica.
5. Estudiemos qué sucedería si el problema fuera más grave. Elimina el 20% de la muestra de individuos tratados al azar y repite el proceso. ¿Qué límites encuentras entonces?