

Ejercicio: Evaluando Métodos Causales para Evaluación de Programas

Fecha de entrega: 1 de junio

Parte 1: Identificando Controles Apropriados

Escenario: Un departamento estatal de educación implementó un programa voluntario de desarrollo profesional docente enfocado en la integración de tecnología en las aulas. Recolectaron datos de 2,000 profesores en todo el estado, incluyendo 500 que eligieron participar en el programa. Quieren evaluar si el programa mejoró los puntajes de las pruebas de los estudiantes.

Las variables disponibles incluyen:

- **experience:** Años de experiencia del profesor
- **highest_degree:** Máximo grado académico del profesor (Licenciatura o Maestría)
- **school_SES:** Nivel socioeconómico promedio de la escuela (: una medida de 10 a 90, más alta significa mayor nivel social)
- **tech_pre:** Competencia tecnológica del profesor medida antes del programa (de 1 a 6)
- **tech_post:** Competencia tecnológica del profesor, medida después del programa
- **student_tech_use:** Nivel de uso de tecnología de los estudiantes en clase, medido al final del año
- **district_type:** Tipo de distrito escolar (urbano, suburbano, rural)
- **test_prev:** Puntajes promedio en pruebas de los estudiantes de ese profesor el año anterior
- **test_curr:** Puntajes promedio en pruebas de los estudiantes de ese profesor en el año actual (post-programa)
- **teacher_motivation:** Motivación del profesor para el crecimiento profesional (encuesta del inicio del año)
- **num_tech_lessons:** Número de lecciones relacionadas con tecnología impartidas (contadas durante el año)

Tu es parte de un time de investigadores encargados de evaluar la eficacia del programa. Sus colegas están discutiendo cómo utilizar la información existente. Para cada afirmación, diga si estás de acuerdo o no y justifique su opinión.

1. "Deberíamos controlar por la competencia tecnológica pre-programa del profesor, ya que los profesores que ya eran expertos en tecnología tenían más probabilidades de inscribirse en el programa y podrían tener mejores resultados estudiantiles independientemente del programa."

2. "Deberíamos controlar por la competencia tecnológica post-programa del profesor para tener en cuenta las diferencias en qué tan bien los profesores aprendieron del programa."
3. "Deberíamos controlar por el nivel de uso de tecnología de los estudiantes en clase, ya que esto refleja qué tan exitosamente el profesor implementó el programa."
4. "Deberíamos controlar por el número de lecciones relacionadas con tecnología impartidas durante el año para aislar el efecto de la calidad más que de la cantidad del programa."
5. "No necesitamos controlar por los años de experiencia del profesor porque es poco probable que esté relacionado con la participación en el programa."
6. "Deberíamos controlar por la motivación del profesor para el crecimiento profesional, ya que esto podría influir tanto en la participación en el programa como en la efectividad de la enseñanza."
7. "Si controlamos por los puntajes pasados, no necesitamos controlar por tipo de distrito o nivel socioeconómico de la escuela, por que la puntaje pasada ya incorpora el efecto de esas variables."

8. Basado en sus respuestas, diseña un diagrama causal para este problema e identifica qué variables deben ser controladas. ¿Existe todavía riesgo de sesgo por variables no observadas? Sí existe, explique y comente al menos una. Si no, explique por qué no.

Parte 2: Evaluando Enfoques Metodológicos

Después de determinar qué variables deberían ser controladas, tres investigadores proponen diferentes enfoques metodológicos. Su tarea es evaluar los argumentos de cada investigador. Cada investigador también respondió con sus preocupaciones sobre cada método propuesto. Responda si las preocupaciones son válidas: si no, explique porque no; sí sí, proponga como se puede lidiar con el problema, o pruebe si es un problema efectivo en la práctica.

Propuesta A: Regresión Lineal

La Dra. Martínez propone: "Deberíamos ejecutar una regresión múltiple de los puntajes actuales de las pruebas sobre la participación en el programa, controlando por todas las variables confusoras que identificamos. Este enfoque es directo y nos permite controlar por muchas variables simultáneamente sin encontrar problemas de dimensionalidad."

A1: La Dra. Martínez añade, "Una ventaja de la regresión es que obtenemos errores estándar precisos que tienen en cuenta toda la variación en nuestros datos." Evalúe este argumento.

A2: El Dr. Wong responde, "Me preocupa que la regresión asuma que los efectos del programa son los mismos para todos los profesores, pero el impacto podría variar según el nivel de experiencia o la competencia tecnológica inicial."

A3: La Dra. Chen añade, "Otro problema es que la regresión asume relaciones lineales entre los controles y los resultados. Si la experiencia del profesor tiene una relación no lineal

con los resultados de los estudiantes, la regresión simple podría producir estimaciones sesgadas."

Propuesta B: Emparejamiento Exacto

El Dr. Wong propone: "Deberíamos usar emparejamiento exacto en todas las variables confusoras. Esto asegura que estamos comparando profesores que son idénticos en dimensiones relevantes sin imponer supuestos de forma funcional como lo hace la regresión."

B1: El Dr. Wong argumenta, "El emparejamiento exacto nos da las estimaciones causales más creíbles porque estamos comparando profesores virtualmente idénticos." Evalúe este argumento.

B2: La Dra. Martínez responde, "Con varias variables continuas como la experiencia del profesor y la competencia tecnológica pre-programa, el emparejamiento exacto nos dejaría con muy pocos emparejamientos, posiblemente ninguno. Además, podríamos tener problemas de soporte común donde algunos tipos de profesores tratados no tienen contrapartes comparables en el grupo de control."

B3: La Dra. Chen pregunta, "Una vez que tenemos los emparejamientos, ¿cómo agregamos exactamente los efectos del tratamiento a través de diferentes grupos emparejados para obtener un efecto general? ¿Y esta estimación representa el ATE o el ATT?"

Propuesta C: Métodos de Puntaje de Propensión

La Dra. Chen propone: "Estimemos puntajes de propensión para la participación en el programa usando regresión logística, y luego usemos ponderación por probabilidad inversa. Esto nos permitirá controlar por muchas variables mientras evitamos los problemas de dimensionalidad del emparejamiento exacto."

C1: La Dra. Chen afirma, "Una gran ventaja de los métodos de puntaje de propensión es que una vez que balanceamos en el puntaje de propensión, hemos balanceado efectivamente todas las variables confusoras."

C2: La Dra. Martínez responde, "Pero los métodos de puntaje de propensión requieren que especifiquemos correctamente el modelo de participación. Si nos perdemos términos de interacción importantes o no linealidades en cómo los profesores se seleccionan en el programa, nuestras estimaciones estarán sesgadas."

C3: El Dr. Wong añade, "Me preocupan los profesores con puntajes de propensión muy altos o muy bajos. Los profesores con una probabilidad del 95% de participar que no participaron recibirán pesos extremadamente altos en IPW, potencialmente distorsionando nuestras estimaciones."

Evaluación Final

Dado el problema de investigación y los datos disponibles:

- ¿Qué método recomendaría y por qué?
- ¿Qué modificaciones haría para abordar las preocupaciones válidas planteadas?
- Si el departamento de educación está particularmente interesado en cómo varía el efecto del programa para profesores con diferentes niveles de experiencia, ¿cómo podría esto influir en su elección metodológica?
- Finalmente, utiliza los datos disponibles [en este enlace](#)¹ para aplicar su método elegido. Presenta su código y su estimación final de los resultados con su interpretación.

¹ Puedes cargar los datos directamente con:

```
datos <- read.csv("http://jmbvgarcia.github.io/assets/code/data_ps3.csv")
```