

# Threats to the internal validity of randomized experiments

Fill In Your Name

30 August, 2021

# Supuestos principales

Deserción

Incumplimiento

Propagación

Efecto Hawthorne

No excluibilidad

# Puntos clave de esta presentación

- ▶ Deserción (variables de interés con datos faltantes)
- ▶ Incumplimiento
- ▶ Propagación
- ▶ Efectos Hawthorne
- ▶ Tratamientos diferenciales en los distintos brazos de tratamiento

## Supuestos principales

# Revisión de los supuestos principales

1. Exclusión
2. No interferencia
3. Asignación aleatoria de sujetos al tratamiento

# Deserción

# Deserción (variables de interés con datos faltantes)

- ▶ Algunas unidades pueden tener datos faltantes para las variables de resultado (= las unidades desertan) cuando:
  - ▶ algunos encuestados no pueden ser encontrados o se niegan a participar en la recolección de datos finales.
  - ▶ se pierden algunos registros.
- ▶ Este es problemático cuando los datos faltantes han sido ocasionados por el tratamiento.
  - ▶ Por ejemplo, las unidades asignadas a control pueden estar menos dispuestas a responder las preguntas de la encuesta.
  - ▶ Por ejemplo, el tratamiento puede haber provocado la migración de unidades y no hay forma de acceder a ellas.
- ▶ Si analizamos los datos descartando unidades con datos faltantes para la variable de resultado, entonces ya estaríamos comparando grupos de tratamiento y control similares.

## ¿Qué podemos hacer?

- ▶ Compruebe si las tasas de deserción son similares en los grupos de tratamiento y control.
- ▶ Compruebe si las covariables de los grupos de tratamiento y control tienen patrones similares.
- ▶ No elimine las observaciones a las que les falten datos para la variable de resultado.
- ▶ A veces, cuando faltan datos para la variable de resultados, podemos **acotar** nuestras estimaciones del efecto del tratamiento.



# ¿Qué podemos hacer?

- ▶ Pero la mejor solución es tratar de anticipar y prevenir la deserción.
  - ▶ Personas que no sepan a qué condición del tratamiento fueron asignadas.
  - ▶ Prometerle al grupo control que va a recibir el tratamiento una vez se termine la investigación.
  - ▶ Planificar ex ante para poder contactar los sujetos en para la recolección final de datos.
  - ▶ Contar con presupuesto para haver seguimiento intensivo a una muestra aleatoria de los sujetos que desierten

# Las covariables con datos faltantes no son tan problemáticas

- ▶ Las **covariables de contexto** (es decir, variables cuyos valores no cambian como resultado del tratamiento) con datos faltantes para algunas observaciones son menos problemáticas.
- ▶ Podemos aprender sobre el efecto causal de un experimento sin esas covariables, como vimos en los módulos [Prueba de hipótesis] ([hypothesistesting.html](#)) y [Estimación] ([estimación.html](#)).
- ▶ También podemos usar la covariable de contexto según lo planeado imputando los datos faltantes.
- ▶ También podemos condicionar nuestro análisis directamente a esos datos faltantes

# Incumplimiento

# Incumplimiento

- ▶ En algunas ocasiones las unidades que son asignadas al tratamiento terminan no recibéndolo. Es decir, no cumplen la asignación.
  - ▶ Cuando todas las unidades asignadas al grupo control no toman el tratamiento, pero sólo algunas unidades asignadas al tratamiento sí lo toman, estamos frente a un caso de incumplimiento unilateral.
- ▶ El efecto de ser asignado al tratamiento no es el mismo que el de recibir el tratamiento.
- ▶ El efecto de recibir el tratamiento es comunmente conocido como “efecto local promedio del tratamiento” (Local Average Treatment Effect, LATE) o efecto promedio del tratamiento para el cumplidor (Complier Average Causal Effect, CACE).
  - ▶ La palabra “local” hace referencia a la idea que el efecto solo ocurre para algunas personas que toman el tratamiento cuando son asignadas a él (el tipo de gente).

- ▶ Se requiere que se cumplan dos supuesto para poder estimar el LATE/CACE a partir de un experimento aleatorio.
  1. Restricción de exclusión: la asignación del tratamiento solo afecta la variable de resultado a través del “recibo” del tratamiento y no a través de ningún otro canal.
  2. Supuesto de monotonicidad: no hay unidades que desafían, es decir, unidades que toman el tratamiento solo cuando son asignadas a control pero no toman el tratamiento cuando son asignados a este.

# Propagación

# Propagación e interferencia entre unidades

- ▶ A veces tendremos sospechas de que la asignación al tratamiento de una unidad afecta los resultados de otras unidades.
- ▶ Si el tratamiento de una unidad afecta el resultado de otra unidad, estamos frente a un caso de violación de uno de los supuestos básicos de la inferencia causal.
- ▶ Podemos seleccionar unidades que están lejos unas de otras para evitar la “transmisión” del tratamiento entre unidades.

# Estudiar los efectos secundarios

- ▶ Esto no es un problema si diseña un estudio que permita investigar la propagación, es decir, que los resultados de una unidad pueden depender del tratamiento de otras unidades.
- ▶ Para investigar los efectos de propagación:
  - ▶ Podemos recolectar datos de la variable de resultado para las unidades que nunca fueron elegibles para la asignación aleatoria al tratamiento, pero que estaban cerca de unidades que sí eran elegibles para el tratamiento y que pudieron verse afectadas por efectos de propagación.
  - ▶ Puede utilizar un diseño de asignación aleatoria en dos etapas.
  - ▶ Puede asignar grupos de unidades (p. Ej., Distritos) a diferentes niveles de intensidad de tratamiento (p. Ej., Usar una proporción diferente de aldeas en los distritos asignados al tratamiento).



# Efecto Hawthorne

# Efecto Hawthorne

- ▶ El efecto Hawthorne: saber que están siendo observados/estudiados puede llevar a los sujetos a comportarse de manera diferente.
- ▶ Esto podría crear un error de medición sistemático tanto en los grupos de tratamiento como en los de control.
- ▶ También podría resultar de una mayor atención prestada al grupo de tratamiento, deshaciendo la creación de grupos de tratamiento y control equivalentes creados a través de la asignación aleatoria

# Buenas practicas

- ▶ Recolectar datos de la forma más discreta posible.
- ▶ En la medida de lo posible, nadie más que el propio sujeto debe conocer a qué condición tratamiento fue asignado.
- ▶ Encuestadores/investigadores deberían desconocer la condición del tratamiento de los sujetos.
- ▶ No haga mediciones adicionales para el grupo de tratamiento.

No excluibilidad

# Diferencias entre los grupos de tratamiento y control más allá del tratamiento

- ▶ Manejar los grupos de tratamiento y control de manera diferente significa que las diferencias observadas en las variable de resultado entre estos grupos pueden deberse al tratamiento y/o al manejo diferente.
- ▶ Los ejemplos incluyen el uso de diferentes cuestionarios para la recolección de datos o rondas adicionales de recolección de datos para el grupo de tratamiento con la intención de obtener información sobre los mecanismos.
- ▶ Recuerde diseñar su estudio y sus encuesta para tratar todos los brazos de tratamiento por igual, excepto por el tratamiento en sí.