Taller 2

Ulises Quevedo

2023-09-24

Taller 2 - Inferencia causal

Suponga que el gobierno está considerando las opciones de dar apoyos monetarios a las escuelas o incentivos monetarios individuales a los profesores con base en el rendimiento que hayan tenido sus alumnos en una prueba estandarizada o los dos juntos. Antes de expandir esta política a todas las escuelas públicas, deciden hacer un piloto y que sea evaluado mediante un experimento aleatorio. Uds. están encargados de diseñaar el experimento. Suponga que Ud. tiene cierto margen de elegir cuantas escuelas pueden estar en este piloto (entre 100 y 200).

1. Diseño del experimento

a) Discuta los méritos de asignar maestros aleatoriamente a que se les den incentivos basados en su rendimiento, sin importar la escuela (i.e., asignar el tratamiento a nivel individual). [Hint: ¿Se cumpliría SUTVA?]

Si bien este puede ser un diseño interesante puesto que aleatoriza completamente la asignación, y maximiza el número de unidades del estudio (cada profesor recibe o no tratamiento, por lo que si hay N profesores tendríamos N unidades), hay algunos problemas respecto a este tratamiento. Principalmente, se viola el SUTVA, puesto que existen externalidades obvias:

En una misma escuela, dos profesores pueden tener los mismos alumnos. Supongamos que un profesor i es asignado al tratamiento: este se esforzará más (pues tiene el incentivo de recibir un beneficio económico) y subirá el puntaje de los alumnos; otro profesor j que no fue asignado a tratamiento, pero tiene los mismos alumnos, también verá el puntaje de sus alumnos subir, incluso si no hizo nada. Luego, Y_j^0 varia debido a T_i, sin importar T_j.

Otra forma en que se podría violar SUTVA es la siguiente: suponga de nuevo que tenemos a i y j, uno asignado y otro no respectivamente. El profesor i podría hablar con j, prometiéndole un porcentaje del apoyo monetario a cambio de que el profesor j también se esuferce para subir los puntajes. Así, Y_i^1 se verá inflado por el hecho de que si bien $T_j = 0$, el compartir el apoyo es como si $T_j = 1$.

b) Discuta los méritos de asignar escuelas aleatoriamente al tratamiento de que a sus maestros se les den incentivos basados en su rendimiento o no (i.e., asignar el tratamiento a nivel escuela). [Hint: ¿Se pierde algo en términos de poder?]

El diseño en el que se incluyen escuelas en el programa de manera aleatoria resuelve los problemas del diseño anterior, puesto que ahora todos los profesores de una misma escuela están o no asignados, y comparten alumnos e incentivos. Sin embargo, este diseño también presenta algunos problemas:

Por un lado, perdemos bastante poder estadístico: en el diseño anterior teníamos N profesores en total. Supongamos que cada escuela tiene, aproximadamente, k profesores. Este nuevo diseño tendría N/k observaciones, puesto que ahora agrupamos a todos los profesores y la unidad de observación es la escuela.

Adicionalmente, podríamos tener el problema (de ser posible) de que un profesor se enterara del programa, y al estar en una escuela asignada al control, decida moverse a una escuela asignada a tratamiento. Este cambio no era posible en el diseño anterior. En ese caso, tendríamos un sesgo de selección que no se elimina con la aleatorización de las escuelas, pues profesores más motivados se moverían al tratamiento si se encuentran en el control.

c) ¿Qué diseño preferiría, y por qué?

Preferiría el segundo diseño, puesto que los problemas que presenta se resuelven más fácilmente: el primero, creciendo el número de escuelas (que si bien es costoso,es resolvible); el segundo, asegurandose de que haya cierta separación entre escuelas o que sea dificil para los profesores hacer el cambio.

El primer diseño, que viola SUTVA completamente, es difícil de arreglar, puesto que tendríamos que garantizar que los alumnos de los profesores tratados y de control no se tocaran para nada, y que los profesores no pudieran coludirse.

d) Supón ahora que el presupuesto que te había otorgado inicialmente el gobierno federal es aumentado y ahora puedes muestrear más escuelas. ¿Cómo afectaría esto en términos de poder?

Si tenemos más escuela, el poder de la regresión aumenta: esto es, es más probable que rechazemos la hipótesis nula dado que es falsa. En términos del experimento, aumentaría la probabilidad de que identifiquemos un efecto si es que en verdad lo hay.

e) Ahora supón que para que la intervención sea considerada por el gobierno para su planeación de política pública te pide que el efecto de la intervención sea mayor o igual a un valor de referencia establecido por la comunidad académica experta en el tema θ, que refleja si una intervención es relevante. ¿Qué parámetro buscarías calibrar en tus cálculos de poder para cumplir con la condición del gobierno federal? ¿De qué otros parámetros depende y de qué forma lo hace?

Para esto, debemos considerar el mínimo efecto detectable (MDE), dada la potencia deseada κ, la significancia deseada α, la proporción de individuos en el grupo de tratamiento p, la variabilidad de la variable de respuesta σ y el número de observaciones N. La relación es:

$$MDE = (t_{\frac{\alpha}{2}} + t_{1-\kappa})\sqrt{\frac{\sigma^2}{p(1-p)N}}$$

Así, primero buscaríamos aumentar N hasta que $MDE = \theta$. Si no se pudiera (pues no hay presupuesto para incluir más escuelas), buscaríamos que $p = \frac{1}{2}$, pues esto maximiza $p(1-p)$. Si incluso así no lo lográramos, tendríamos que encontrar un compromiso, sea reduciendo la potencia (estamos dispuestos a reducir la probabilidad de rechazar cuando sí hay un efecto) o aumentando la significancia (estamos dispuestos a aumentar la probabilidad de rechazar cuando no hay efecto).

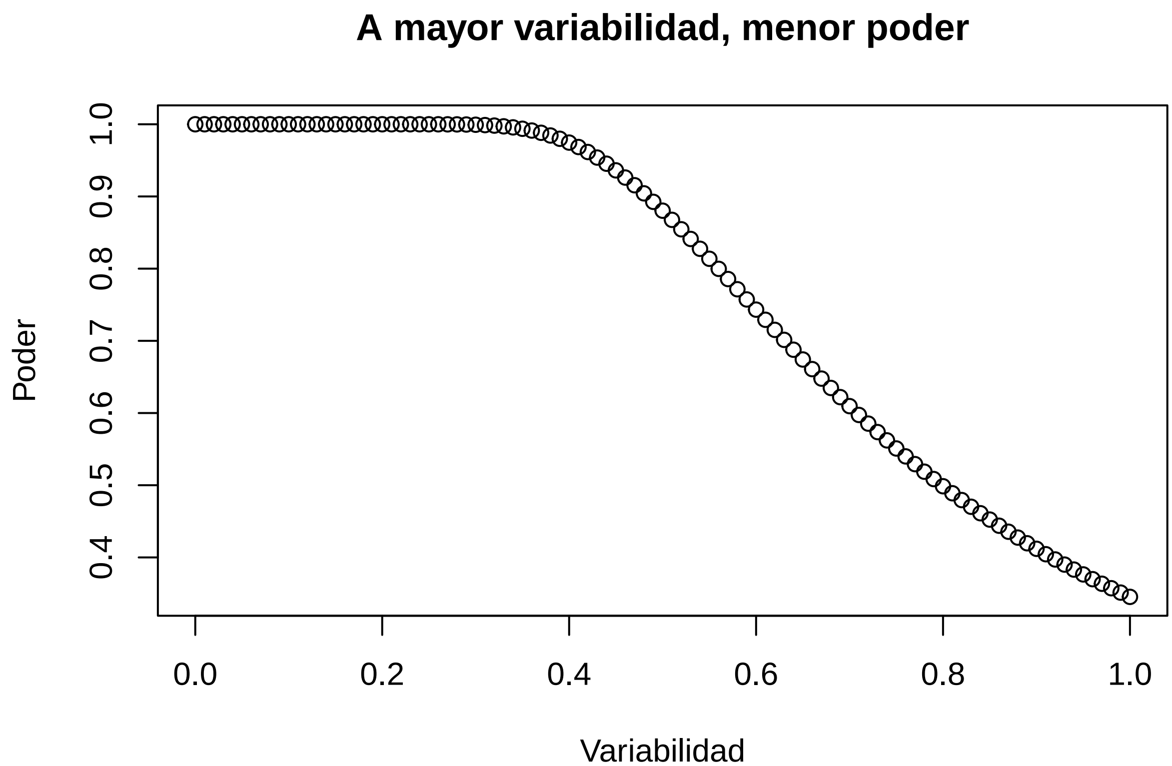
f) En la evaluación de impacto de esta intervención consideras algunas variables de respuesta para evaluar el efecto del tratamiento. ¿Cómo afecta una alta varianza poblacional de la variable de respuesta en el poder estadístico del experimento?

En primer lugar, regresamos a la ecuación de arriba y observamos que a mayor varianza, más alto es el MDE; intuitivamente, esto nos dice que a mayor variabilidad, menor poder estadístico, pues necesitamos efectos más grandes para poder detectar (y rechazar la nula) con éxito.

De la misma ecuación de arriba, podemos fijar todo excepto κ y σ; luego, vemos que a mayor σ, si queremos seguir detectando el mismo efecto θ, tendremos una menor potencia.

Observemos: supongamos que tenemos un efecto que deseamos detectar $\theta = 0.1$, que tenemos $p = 0.5$, $\alpha = 0.05$,$N = 1000$. Ahora, veamos la relación entre el poder que podemos obtener manteniendo esto fijo, a medida que aumenta la variabilidad. Recordemos que la relación de κ en función de sigma es (con P la función de distribución de una t-student con df grados de libertad:

$$\kappa = 1 - P(t \geq \frac{\beta}{\sqrt{\frac{\sigma^2}{Np(1-p)}}} - t_{\frac{\alpha}{2}})$$



g) ¿De qué manera tendrías que establecer el tamaño de los grupos de tratamiento y control para maximizar el poder estadístico, todo lo demás constante?

Notemos que maximizar el poder estadístico (todo lo demás constante) es equivalente a minizmiar el Efecto Mínimo Detectable. Luego, queremos minimizar la expresión (k) respecto a p. Notemos que p aparece en el denominador de la raíz. Luego, la raíz se minimiza cuando $p(1-p)$ se maximiza, lo cual ocurre cuando $p = \frac{1}{2}$