

Microeconometría aplicada

Taller 2

Mauricio Romero

10 de octubre de 2020

1. Instrucciones generales

- Pueden usar R, Stata o cualquier otro programa
- Deben entregar un archivo PDF. Algo así como un memo con los que hicieron y lo que encontraron.
- Adicionalmente, deben entregar un archivo ZIP con los datos crudos (raw data), y todos los códigos (scripts) que usen.
- Yo debería poder replicar sus resultados descomprimiendo el archivo zip y cambiando (únicamente) el directorio principal al principio y corriendo el código.
- Envíen sus respuestas (PDF y ZIP) a: mtromero+microeconometria@itam.mx
- Van a ser evaluados no solo en la respuesta final. También tendré en cuenta el proceso que usen para llegar a sus respuestas (i.e., su código y los comentarios de ese código).
- Es muy importante que escriban su código (y los comentarios) de manera que sea fácil para otros (yo en este caso) seguir el código y entender lo que se hizo.
- Usar recursos adicionales (internet, libros, blogs, etc.) está bien (y muchas veces es necesario), pero por favor menciones los recursos que usan y cítelos adecuadamente.
- La fecha de entrega es octubre 15 de 2020 a las 12:00 (medio día)
- La tarea se puede hacer en grupos (del tamaño que quieran)

2. Problema 1 - Experimento en educación

Suponga que el gobierno está considerando una política donde se le entrega la administración de escuelas públicas a entidades privadas (e.g., organizaciones de la sociedad civil, entidades sin ánimo de lucro enfocadas en educación, universidades privadas, o entidades con ánimo de lucro enfocadas en educación). 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 diseñar el experimento. Suponga que Ud. tiene cierto margen de elegir cuantas escuelas pueden estar en este piloto (entre 100 y 200). Suponga que 4 organizaciones privadas van a participar en el piloto, y cada una va a manejar el mismo número de escuelas.

1. Primero debemos diseñar el experimento

- a) Discuta los méritos de asignar escuelas aleatoriamente a diferentes organizaciones (i.e., asignar cada escuela a uno de 4 posibles organizaciones o al control).

- b) Discuta los méritos de primero separar la muestra en 4 grupos (basado en las preferencias geográficas de cada organización) y después asignar las escuelas a tratamiento o control dentro de cada grupo.
- c) ¿Qué diseño preferiría, y por qué?
- Suponga que se hizo el primer diseño (i.e., se asignaron escuelas aleatoriamente a uno de 4 posibles organizaciones o al control). Antes de hacer cualquier análisis, ¿Ud. cree que comparar los resultados de los estudiantes en las escuelas de tratamiento con las de control unos años después del tratamiento nos va a indicar la efectividad de esta política pública? [Pista: Piense en la decisión de los padres de a que escuela deciden inscribir a los niños, y en los incentivos de las organizaciones privadas a inscribir/desinscribir cierto tipo de estudiantes]
 - Suponga que efectivamente comparando los resultados de los estudiantes en diferentes escuelas puede estudiar el efecto del tratamiento. Suponga que Ud. tiene una base de datos que tiene resultados a nivel estudiante para diferentes escuelas: ¿Que regresión utilizaría para medir el efecto del tratamiento (indique el nivel de cada variable, como manejaría los errores estándar, si va a incluir controles y cuales, etc.)?
 - Haga cálculos de poder para este experimento usando mínimos cuadrados ordinarios para analizar los datos. Es decir, simule diferentes efectos del tratamiento (incluido cero) y muestre con que probabilidad rechazaría la hipótesis nula. Para esto, utilice los datos de PLANEA para grado Sexto de 2017-2018 en matemáticas a nivel estudiantes, los cuales puede encontrar aquí https://www.inee.edu.mx/wp-content/uploads/2019/01/Planea06_2018_Alumnos.zip (más información aquí <https://www.inee.edu.mx/evaluaciones/planea/sexta-primaria-ciclo-2017-2018/>). Para ser más específico. Elija una muestra de 200 escuelas públicas, asigne las escuelas aleatoriamente a tratamiento y control, y después simule diferentes efectos de tratamiento.
 - Repita los cálculos de poder, utilizando inferencia de aleatorización
 - Un experimento muy similar a este se llevó a cabo en Liberia (ver <http://mauricio-romero.com/pdfs/papers/aer.20181478.pdf>). Los datos de ese experimento están en: <http://mauricio-romero.com/pdfs/FieldExperiments/Bern2020/Data/Student.csv>. Las variables indican (en este orden): Si la escuela fue tratada o no, el id de la escuela, una variable que no es relevante, el género del estudiante, el peso muestral de la observación, el grado del alumno, el puntaje en inglés (estandarizado para tener media cero y desviación estándar de uno), el puntaje en matemáticas (estandarizado para tener media cero y desviación estándar de uno), y la edad del estudiante.
 - Estime el efecto del tratamiento (y el p-valor) en inglés y matemáticas usando mínimos cuadrados ordinarios.
 - Estime el efecto del tratamiento (y el p-valor) en inglés y matemáticas usando inferencia de aleatorización.

3. Problema 2 - Experimentos con múltiples tratamientos

Considere el siguiente diseño experimental:

		T_1	
		No	Si
T_2	No	N_c	N_T
	Si	N_T	0

Donde se aleatorizan dos tratamientos (T_1 y T_2). N_c y N_T denotan el número de observaciones que se asignan a cada grupo (y $N = N_c + 2 * N_T$ esta fijo). Suponga que este experimento se analiza con el siguiente modelo de regresión:

$$Y = \beta_0^* + \beta_1 T_1 + \beta_2 T_2 + \varepsilon. \quad (1)$$

1. Sea $\hat{\beta}_1$ el estimador de mínimos cuadrados ordinarios. ¿Cuál es la varianza de $\hat{\beta}_1$?
2. ¿Cuál es la asignación óptima de la muestra que maximiza el poder del t-estadístico para $\hat{\beta}_1$? En otras palabras, ¿Cuál es la asignación de la muestra total que minimiza la varianza de $\hat{\beta}_1$?

4. Problema 3 - ENOE

Utilizando la base de datos de la ENOE 2020 - Primer trimestre, realice los siguientes ejercicios. La base y la documentación asociada la puede encontrar aquí: <https://www.inegi.org.mx/programas/enoe/15ymas/>

1. Estime la desigualdad del ingreso según la ENOE (GINI, percentiles del ingreso alto, y con qué % del ingreso se queda la gente más rica)
2. ¿Es diferente de la que habíamos encontrado con la base del SAT? ¿Qué puede explicar la diferencia?
3. Estime el efecto de la educación en el nivel de ingresos. Presente sus resultados con una tabla de regresión que sea fácil de leer y estética. Interprete los resultados de las variables: genero, educación (en años), y edad. Controle por otras variables que considere pertinentes (justifique por que las incluyo y como). ¿Es esta estimación causal?
4. Haga esta regresión para los hombres y para las mujeres por aparte. ¿Como se relacionan los coeficientes de estas dos regresiones, con la que usa los datos de toda la población?
5. Estime el retorno de la educación en el logaritmo de los ingresos. Presente sus resultados con una tabla de regresión que sea fácil de leer y estética. Interprete los resultados de las variables: genero, educación (en años), y edad. Controle por otras variables que considere pertinentes (justifique por que las incluyo y como). ¿Es esta estimación causal? ¿Hay algún problema al tratar de usar el logaritmo del ingreso?