

Microeconometría aplicada

Taller 1

Mauricio Romero

24 de agosto 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 cítenlos adecuadamente.
- La fecha de entrega es septiembre 11 de 2020 a las 12:00 (medio día)
- La tarea se puede hacer en grupos (del tamaño que quieran)

2. Problema 1 - Ecobici

Bajen los datos de Ecobici de lo que va corrido de 2020 (<https://www.ecobici.cdmx.gob.mx/es/informacion-del-servicio/open-data>)

- Calcule el número promedio de viajes diarios y haga una gráfica que muestre su evolución. Ponga en la gráfica eventos importantes en lo corrido del año (e.g., cuando empezó la jornada de sana distancia, cuando acabo, cuando la ciudad cambio de semáforo rojo a naranja, y otros que considere importantes)
- Haga un histograma de los viajes diarios. Haga otra gráfica con un histograma (o un density plot) para cada día de la semana
- Calcule la distancia total recorrida por los usuarios cada día y haga una gráfica que muestre su evolución
- Hagan una gráfica que muestre cuales son las bicicletas más rápidas y cuáles son las más usadas. Realice una gráfica que muestre si hay alguna relación entre “más rápida” y “más usada”

- Haga una visualización de donde a donde hacen más viajes las personas esto. Tenga en cuenta la hora del día. La mejor visualización recibirá puntos extra en el primer examen

3. Problema 2 - Desigualdad del ingreso

Descarguen los datos anonimizados de las declaraciones de personas físicas de 2015 <https://datos.gob.mx/busca/dataset/informes-del-articulo-decimo-noveno-transitorio-de-la-ley-de-ingresos-de-la-federacion-lif-2017>: Anuales ISR Personas Físicas 2015 El link del diccionario de datos no funcionaba el 24-Agosto-2020, pero lo pueden encontrar acá: https://www.dropbox.com/s/l3j0iw32b1z472v/DiccionarioDatosPF_anonimizados_2010-2015.xls?dl=0

- Calcule el coeficiente de GINI (la variable adecuada es “TOTAL DE INGRESOS ACUMULABLES”) y haga una visualización del cómo se hace este calculo
- Haga un histograma con los ingresos
- Calcule con qué porcentaje del ingreso se queda el: 0.1 % con mayores ingresos, el 1 % con mayores ingresos, el 5 % con mayores ingresos y el 10 % con mayores ingresos. ¿Cree que estos numero subestiman o sobreestiman el valor real? Explique su respuesta
- Haga una visualización sobre la desigualdad del ingreso. La mejor visualización recibirá puntos extra en el primer examen

4. Problema 3 - Desigualdad de género

Descargue los datos de ingresos de los servidores públicos que trabajan en la Secretaría de Gobierno de CDMX de 2019 <https://www.transparencia.cdmx.gob.mx/secretaria-de-gobierno/entrada/372>

- Calcule el ingreso anual de los empleados y haga una tabla con las estadísticas descriptivas del ingreso anual
- Haga una gráficas donde muestre la distribución del ingreso (histogramas o density plots) por genero
- ¿Es la diferencia en salarios entre hombres y mujeres estadísticamente significativa?
- ¿Qué factores en los datos pueden explicar esta diferencia? Realice un análisis visual
- ¿Estos factores relevan más o menos discriminación de género? Explique su respuesta

5. Problema 4 - Teorema del límite central

Haga el siguiente ejercicio con para la siguientes distribuciones: $N(7, 8)$, $\Xi^2(3)$, $Exp(5)$, $U(0, 1)$

- Simule 10000 observaciones de la variable aleatoria y calcule la media
- Haga esto 1, 10, 100, 1000 veces y grafique la distribución de la media (un histograma)
- Sobreponga la distribución teórica según el teorema del límite central
- Haga una visualización sobre el teorema del límite central. La mejor visualización recibirá puntos extra en el primer examen

6. Problema 5 - Ley de los grandes números

Haga un programa que simule una ruleta (de casino). Suponga que Ud. tiene 1,000 pesos que puede apostar como quiera.

- Proponga tres estrategias de apuesta y simúelas. Digamos que máximo puede jugar X veces (después de eso cierran el casino). Calcule la ganancia esperada (usando simulaciones). Haga esto para $X=1, 10, 100, 1000$
- Calcule la ganancia esperada (teórica) de esas estrategias