

TEORÍA ECONOMETRICA II

Tarea 1

1ER SEMESTRE 2023

Profesor: Tomás Rau
Ayudantes: Eduardo Barrueto y Nicolás Valle

En esta tarea Ud. debe entregar via Canvas, hasta el 5 de abril, 23:59, un reporte (preferentemente realizado en Latex) con las respuestas a cada una de las preguntas, como también los códigos computacionales que permitan replicar sus estimaciones. Respuestas sin códigos tendrán calificación mínima. Les recuerdo que las tareas son individuales y que los atrasos en la entrega de las tareas serán penalizados con 5 décimas por día de atraso.

Pregunta 1: Modelos Logísticos (60 puntos)

En el presente ejercicio implementaremos algunos modelos logísticos para analizar la decisión que realizan los apoderados sobre el colegio al que postulan a sus hijas/os.

La base de datos `data_pregunta1.dta` cuenta con información sobre las preferencias declaradas por una muestra de apoderados de la Región Metropolitana en el periodo principal del SAE. La base posee las siguientes variables: 1) características del colegio al cual postula (`tipo`, `convenio_pie`, `municipal`), 2) características del postulante como sexo (`es_mujer`), variables binarias si tiene hermano en el colegio que postula (`prioridad_hermano`), si es considerado como alumno prioritario (`prioritario`) o alto rendimiento (`alto_rendimiento`) y al nivel al que postula (`cod_nivel`). Además, se reportan la distancia (`distanciaX`), el puntaje `simce` de matemáticas y lenguaje (`simce_mateX` y `simce_lectX`) y el porcentaje de alumnos prioritarios (`p_prioritarioX`) de sus otras opciones.

- Considere solamente el tipo de administración del establecimiento seleccionado (si es municipal o no). Programe en STATA la función de verosimilitud para un modelo Logit, luego, estímelo para la decisión de postular a un colegio municipal. Como variables explicativas considere las características del postulante (i.e., curso al que postula (`cod_nivel`), sexo (`es_mujer`), si es considerado estudiante prioritario¹) o alto rendimiento². Sea explícito en el algoritmo de optimización, la tolerancia y el tipo de errores estándar que ejecuta en su rutina estadística. Justifique sus decisiones con argumentos estadísticos. **(10 puntos)**
- Obtener los Efectos Marginales para cambios en las características del individuo e interprete sus resultados. **(5 puntos)**
- Los establecimiento educacionales además de diferenciarse por tipo de administración se pueden clasificar si tienen programa de integración o no (`convenio_pie`). Genere los labels para cada tipo especificando primero el tipo de administración y luego si tiene convenio pie. Tabule cuántos colegios hay de cada tipo (1 al 4). ¿Cuál sería la principal desventaja de usar el tipo de administración solamente (sólo usar la variable `municipal`)? **(5 puntos)**

¹Los/as alumnos/as prioritarios/as son aquellos para quienes la situación socioeconómica de sus hogares puede dificultar sus posibilidades de enfrentar el proceso educativo. En el siguiente link puede revisar más información <https://www.ayudamineduc.cl/ficha/alumnos-prioritarios-2022>

²Se considera que los estudiantes tienen un rendimiento académico alto si provienen del 20% superior del ranking de calificaciones de su escuela.

- d. Considere ahora las cuatro alternativas y estime un Multinomial Logit utilizando nuevamente las características del individuo. Obtenga los efectos marginales. Interprete sus resultados y compárelos con los obtenidos en el acápite anterior. **(10 puntos)**
- e. Su prima le recomienda usar las características (distancia, puntajes Simce y porcentaje de alumnos prioritarios) de las 4 alternativas disponibles (variable tipo³) para analizar la decisión tomada por el individuo. Desde la teoría económica, piense en las ventajas o desventajas que tendría esta especificación. Luego, estime un Conditional Logit siguiendo la recomendación. Obtenga los efectos marginales. Interprete sus resultados. **(10 puntos)**
- f. Existe un tipo de modelo mixto que permite estimar elección discreta con ambos tipos de regresores: que varían con la alternativa y aquellos que no pero varían con la persona. Este se debe a McFadden (1974) y algunos autores le llaman Mixed Logit o Alternative Specific Conditional Logit. Estime este modelo considerando toda la información a su disposición. Obtenga los efectos marginales. Interprete sus resultados y compárelos con los obtenidos en el acápite anterior. **(15 puntos)**
- g. Otra prima le comenta que en el SAE los apoderados rankean los colegios, por lo que podría utilizar un Ordered Logit. Explique el modelo que le plantea su otra prima y si es posible implementarlo con la información disponible. Si la respuesta es si, obtenga los efectos marginales y si es no, explique detalladamente qué información falta y cómo debería ser la estructura de los datos. **(5 puntos)**

Pregunta 2: Probit binario y multinomial (60 puntos)

En base a la lectura de “Names and Reputations: An empirical analysis”:

- a. Resuma brevemente el artículo con énfasis en las predicciones teóricas y en como los resultados se alinean con las mismas. **(10 puntos)**
- b. Usando la base AEJMicro-2009-0153_data.dta, replique la tabla 5 e interprete sus resultado ¿Se reportan los efectos marginales evaluados en la media de los regresores o un promedio de los efectos marginales de cada individuo? Incluya una en su tabla una columna que indique los efectos marginales en la versión alternativa a la expuesta en la tabla del artículo ¿Qué puede decir de las diferencias? **(15 puntos)**
- c. Replique la tabla 6 e interprete sus resultados. En sus interpretaciones sea preciso al referirse al rol de la alternativa *base* ¿Cuál es la necesidad econométrica de que exista esa alternativa? **(15 puntos)**
- d. Vuelva a replicar las tablas 5 y 6 pero esta vez utilice un modelo logit. Compare sus resultados ¿Es posible comparar directamente ambas estimaciones? Justifique. De ser posible corregir los resultados logit para poder comparar, hágalo y luego interprete. **(10 puntos)**
- e. Replique las tablas 7 y 8. Interprete sus resultados y refiérase a las hipótesis que cada tabla busca poner a prueba. **(10 puntos)**

Pregunta 3: Modelos de Conteo (50 puntos)

A partir de la lectura “Victims Incentives and Criminal Activity: Evidence from Bus Drivers Robberies in Chile” responda las siguientes preguntas:

- a. Explique en qué contexto y con qué fin se hace uso de un modelo de conteo en este estudio. Además, explique la estrategia empírica y sus supuestos. Extensión máxima: 1 página. **(10 Puntos)**
- b. Utilizando la base de datos entregada para la tarea (panel_muni_week.dta) replique los resultados de la Tabla 2. Interprete sus resultados. **(10 puntos)**

³La variable tipo toma el valor de 1 si el establecimiento no es municipal y no tiene convenio pie; el valor de 2 si el establecimiento es municipal y no tiene convenio pie; el valor de 3 si el establecimiento no es municipal y posee convenio pie; y el valor de 4 si es municipal y posee convenio pie.

- c. Ud. se da cuenta que la variable de interés es una variable de conteo. Vuelva a realizar la Tabla 2 pero utilizando un modelo de conteo (poisson y binomial negativa). Interprete sus resultados. Compare sus resultados con los encontrados en el inciso anterior. **(20 puntos)**
- d. ¿Existen diferencias en los resultados? ¿Por qué es utilizar un modelo de conteo es una buena aproximación para este tipo de estimaciones? **(10 puntos)**