

## Tarea 3

**Fecha de entrega: Lunes 7 de diciembre**

**Nota:** Deberán subir a *Canvas* un archivo de texto con sus respuestas. Pueden utilizar el formato de su preferencia (e.g. L<sup>A</sup>T<sub>E</sub>X, Word u hojas escritas a mano y escaneadas). Además, deberán subir otro archivo que genere todos sus resultados de las preguntas prácticas. Puede ser un R-script, Do-File o script de algún otro software similar.

En preparación para su examen deberán llevar a cabo un análisis sencillo utilizando el modelo de Logit Multinomial. Este análisis es individual, lo cual quiere decir que no deben consultar con sus compañeros la forma de llevar a cabo esta tarea. [Tip: en Stata el logit multinomial se hace con el comando `mlogit` y en R con la librería `mlogit`]

Para este ejercicio utilizarán la base de datos *Stove.dta* que incluye información acerca del tipo de estufa que diferentes hogares en San Diego han instalado. Las posibles alternativas de estufa que los hogares podrían haber elegido son: *gc* para gas central, *gr* para gas individual, *ec* para eléctrica central, *er* para eléctrica individual, y *hp* que representa otro tipo.<sup>1</sup> Las variables explicativas disponibles incluyen datos del hogar, como lo son: *income*, que es una variable categórica de 2 a 7 para el nivel de ingreso (donde 7 es mayor ingreso), *age* para la edad del jefe del hogar, *num\_people* para el número de personas que habitan el hogar, *region* que describe la región donde habita el hogar, *ic.estufa* que son 5 variables que indican el costo de instalación de una estufa de cada tipo<sup>2</sup> y *oc.estufa* que son 5 variables que indican el costo anual (en USD) de operación de cada tipo de estufa para cada hogar.

1. Iniciemos explorando la base de datos. Elabora una tabla con estadísticas descriptivas básicas para las variables de la base de datos. Tu deberás de decidir cuáles son las variables relevantes a incluir.
2. En tu base de datos existen 4 posibles regiones: valley, scostl, ncostl y mountn. Haz una tabla en la que indiques para cada región el porcentaje de hogares que han adquirido cada tipo de estufa. Es decir, los porcentajes de cada región deben sumar 100. Los renglones de tu tabla deben ser el tipo de estufa y las columnas las distintas regiones.
3. Haciendo una estimación de logit multinomial muestra cómo podrías replicar esta tabla. Debes reportar el logit multinomial que hayas estimado en una tabla de regresión con el formato típico que hemos usado durante el semestre. Usando los resultados

---

<sup>1</sup>Las estufas individuales son de casas particulares y las *central* son aquellas que a veces se instalan en condominios de forma conjunta igual a todos los hogares.

<sup>2</sup>Por ejemplo, la variable *ic.gr* indica el costo en dólares de instalación de una estufa de gas individual. Esta es una variable que cambia por tipo de estufa y por hogar, ya que los precios de instalación se negocian individualmente. El caso de las variables *oc.estufa* son similares.

muestra como puedes replicar el resultado del porcentaje de hogares en la región *mountn* que adquieren una estufa gr.

4. Lleva a cabo una estimación de logit multinomial ahora incluyendo todas las variables explicativas que describimos arriba. Reporta tus resultados en una tabla. Tu deberás decidir cuándo usar dummies, y cuando variables por si solas. Es importante considerar que para el costo de instalación y de operación deberás obtener un solo coeficiente. Para las demás variables si deberás obtener un coeficiente para cada alternativa de estufa.
5. Utilizando los resultados de la tabla anterior deberás calcular el *efecto parcial promedio* para la probabilidad de adquirir una estufa *er* por aumentar en dos personas el número de habitantes en el hogar. Indica la interpretación más específica posible para este resultado.
6. (Puntos extra) En MCO utilizamos la  $R^2$  para evaluar qué tan buenas son las predicciones del modelo. Sugiere una forma de evaluar qué tan buenas son las predicciones de tu modelo de logit multinomial y reporta los resultados de esta evaluación. Tip: Piensa que un *error* en el contexto del modelo del logit multinomial puede ser una clasificación incorrecta de un hogar.