

Profesor: Héctor Bahamonde, PhD.

e: hector.bahamonde@uoh.cl

w: www.HectorBahamonde.com

Curso: MLE.

TA: Gonzalo Barriá.

INSTRUCCIONES GENERALES

Usando la base de datos que será asignada al azar y dentro de los plazos establecidos en la sección *Tareas* de uCampus, deberás:

1. **Generar una pregunta de investigación.** Un ejemplo de pregunta de investigación puede ser *Qué explica Y?* Para esto, ayúdate del *codebook* (archivo de texto que contiene el nombre de la variable y su definición).
2. **Generar una hipótesis de trabajo.** Una hipótesis de trabajo es una respuesta (una afirmación) *hipotética* a esta pregunta. Un ejemplo podría ser: *Y es explicado por X_1* . En general, aquí también explicas *por qué* X_1 explica *Y*.
3. **Seleccionar las especificaciones que más estimes convenientes.** Para ello deberás estimar GLM's usando técnicas de MLE. Tú deberás justificar la especificación que más estimas conveniente. **Ten en cuenta que si fallas en este paso, todo el resto del trabajo estará malo. Te sugiero que te juntes conmigo en horario de atención de alumnos.** En esta sección deberás **usar todos los tests necesarios** para justificar que tu especificación está correcta. **Será de tu exclusiva responsabilidad testear todos los supuestos distribucionales abordados en clases.** Recuerda incluir **todos** los diagnósticos pertinentes (remítete a la clase “*Diagnósticos*”).
4. **Deberás tener controles.** Recuerda que los controles son otras variables que entran en la misma ecuación, y que ayudan a mantener los efectos constantes en las medias de esas variables. Para cada control deberás justificar la decisión de incluir los controles que incluiste. Esto es discursivo, y como vimos en clases, deriva de la teoría, la revisión bibliográfica, y hasta de la intuición.
5. **Testear la hipótesis.** *Testea* tu hipótesis de trabajo. Recuerda entonces que deberás formular una pregunta de investigación y una hipótesis de trabajo que puedas testear. Para efectos del trabajo, no te servirá formular una hipótesis (o pregunta), si es que no tienes variables relevantes para contestar la pregunta de investigación. Entonces la recomendación es que orientes tus hipótesis y pregunta de investigación a la información que tengas disponible en la base de datos.
6. **Interpretación y presentación de resultados.** Usa **todas** las técnicas específicas de cada modelo para su interpretación. Concentra tus esfuerzos en lo substantivo a la hora de “construir perfiles” (si es que fuera necesario). Deberás también interpretar los controles.
7. **Presentación final en clases.** En la presentación final deberán abordar todos estos puntos. La presentación no deberá durar más de quince minutos. *Todos los integrantes del grupo deberán presentar.* La idea de presentar es que puedan recibir comentarios del resto de los compañeros—estas preguntas cuentan como participación. El trabajo del grupo será convencer a la audiencia que han hecho todo lo posible para (1) encontrar la especificación correcta y (2) *testear* la hipótesis de la mejor manera posible.
8. **Productos finales:**

- (a) **Trabajo Final:** un *script* de R enviado por uCampus. No habrá un informe final. Asegúrate que el *script* corra sin problemas, que los paquetes que usas estén declarados, etc.
- (b) **Presentación final:** un *Power Point* presentado en clases por todos los integrantes del grupo. Todos los aspectos señalados arriba deberán estar abordados en esta presentación.

BASES DE DATOS

Este semestre asignaremos al azar alguna de las siguientes bases de datos.

Librería	Base de Datos	Codebook	Variable Dependiente
library(AER)	data(Fatalities)	help(Fatalities)	Fatalities\$fatal
	data(Guns)	help(Guns)	Guns\$law
	data(PhDPublications)	help(PhDPublications)	PhDPublications\$articles
	data(ResumeNames)	help(ResumeNames)	ResumeNames\$call
	data(TravelMode)	help(TravelMode)	TravelMode\$mode
	data(Arrests)	help(Arrests)	Arrests\$released
library(carData)	data(Chile)	help(Chile)	Chile\$vote