

Maestría en Ciencia Social con especialidad en Sociología
Centro de Estudios Sociológicos
El Colegio de México
Semestre: Enero – Junio 2020
Programa del curso: Análisis de datos categóricos
Docente: Patricio Solís

Introducción

Es muy frecuente que en el análisis sociológico de métodos cuantitativos encontremos variables dependientes de tipo nominal o “categóricas”. Para este tipo de situaciones existen un conjunto de técnicas que se agrupan bajo el nombre de “Análisis de datos categóricos”. Esto incluye a las técnicas descriptivas o exploratorias, como el análisis de contingencia, así como los métodos que permiten extender las técnicas de regresión tradicionales a situaciones en las que la variable dependiente es de tipo categórico. Entre estas técnicas se encuentran la regresión logística binomial y multinomial. Asimismo, existen métodos como la regresión loglineal, que nos permiten modelar el patrón de asociación entre dos o más variables categóricas.

Las técnicas referidas forman parte de un arsenal de formación metodológica básico para cualquier investigador en ciencias sociales, por lo que su aprendizaje debe seguir inmediatamente a la formación básica en regresión lineal.

En el curso revisaremos estas técnicas, con énfasis en sus fundamentos lógicos y la práctica de su aplicación. Esto implica aprender a reconocer sus supuestos básicos, sus alcances y sus limitaciones. También demanda la adquisición de habilidades para aplicarlas en situaciones reales de investigación.

La dinámica del curso consistirá en la exposición en clase de los fundamentos teóricos y prácticos de cada técnica, seguida por ejercicios prácticos, en los que los estudiantes podrán ensayar directamente la aplicación de las técnicas mediante el uso de software especializado (Stata).

Para ingresar a este curso los alumnos deberán poseer conocimientos básicos de estadística inferencial, así como manejar la técnica de regresión lineal.

Objetivo

Al finalizar el curso los alumnos serán capaces de interpretar y utilizar los modelos de regresión logística binomial, logística multinomial, logística ordenada, y log-lineal, en situaciones reales de investigación, así como identificar y resolver los problemas técnicos más comunes asociados a la aplicación de estas técnicas.

Evaluación

La evaluación del curso se realizará de la manera siguiente:

10% Asistencia y participación en clase.

20%	Primer examen parcial
25%	Segundo examen parcial
10%	Propuesta de trabajo final
35%	Trabajo final

Descripción de sesiones*

Sesión	Sesión	Clases	Sesiones de laboratorio (1.5 horas)
1	Sesión 1	Introducción al curso. Revisión del análisis de contingencia en tablas de doble entrada. Pruebas de asociación en el análisis de contingencia. Riesgos. Riesgos relativos. Momios. Razones de momios.	Ejercicio en clase: Tablas de contingencia. Cálculos manuales de riesgos, riesgos relativos, momios, razones de momios. Pruebas de asociación estadística.
2	Sesión 2	El modelo lineal de probabilidad y sus limitaciones. El modelo lineal generalizado. Introducción al modelo de regresión logística binomial.	Réplica de las notas de clase en Stata
3	Sesión 3	Interpretación de coeficientes y razones de momios en el modelo de regresión logística binomial. Probabilidades estimadas	Réplica de sesión 3 en Stata, dudas prácticas sobre uso de Stata para cálculo de coeficientes, razones de momios y probabilidades estimadas
4	Sesión 4	Estimación por máxima verosimilitud del modelo de regresión logística binomial. Medidas de bondad de ajuste. Pseudo R cuadrada de McFadden. Prueba de Hosmer y Lemeshow. BIC.	
5	Sesión 5	Ejercicio en clase: ajuste de modelos de regresión logística binomial. Interpretación de coeficientes. Evaluaciones de bondad de ajuste	Revisión de ejercicio en clase
6	Sesión 6	Primer examen parcial	
7	Sesión 7	Regresión logística multinomial. Introducción. Presentación formal del modelo. Selección de categoría base de contraste. Supuesto de Independencia de Alternativas Irrelevantes (IIA)	
8	Sesión 8	Interpretación de coeficientes, probabilidades ajustadas y bondad de ajuste en el modelo de regresión logística multinomial	Revisión de primer examen parcial
9	Sesión 9	Ejercicio en clase: ajuste de modelos de regresión logística multinomial. Interpretación de coeficientes. Evaluaciones de bondad de ajuste	
10	Sesión 10	Regresión logística ordenada. Presentación formal del modelo. Supuesto de Momios Proporcionales. Interpretación de coeficientes. Evaluaciones de bondad de ajuste	Revisión del ejercicio en clase sesión 9
11	Sesión 11	Ejercicio en clase: ajuste de modelos de regresión logística ordenada. Interpretación de coeficientes. Evaluaciones de bondad de ajuste	Revisión del ejercicio en clase sesión 11
12	Sesión 12	Segundo examen parcial	
13	Sesión 13	Revisión del examen parcial	Revisión del examen parcial
14	Sesión 14	Discusión de propuestas de trabajo final	
15	Sesión 15	Uso de ponderadores en encuestas	
16	Sesión 16	Sesión de aclaración de dudas para trabajo final	

* Es posible que el contenido de las sesiones cambie en función de los avances logrados por el grupo a lo largo del curso.

Materiales de referencia básicos:

Agresti, Alan. 2007. An introduction to Categorical Data Analysis. (2ª Edición). New Jersey: John Wiley & Sons.

Long, Scott y Freese, Jeremy. 2006. Regression Models for Categorical Dependent Variables Using Stata. College Station: Stata Press

Solís, Patricio. 2016. "Aspectos metodológicos en el análisis de la movilidad social" En Solís, Patricio y Marcelo Boado (coords.): Y sin embargo se mueve...: Estratificación social y movilidad intergeneracional de clase en América Latina. México: El Colegio de México - Centro de Estudios Espinosa Yglesias.

Otra bibliografía de consulta:

Hosmer, David H. y Stanley Lemeshow. 2000. Applied Logistic Regression. (2ª Edición). New Jersey: John Wiley & Sons

Powers, Daniel A. & Yu Xie 2000. Statistical methods for categorical data analysis. New York: Academic Press

Treiman, D. J. 2009. Quantitative Data Analysis: Doing Social Research to Test Ideas. John Wiley & Sons

Recursos “en línea”:

Regresión logística binomial:

<http://www.ats.ucla.edu/stat/stata/dae/logit.htm>

Regresión logística ordenada:

http://www.ats.ucla.edu/stat/stata/output/stata_ologit_output.htm

Regresión logística multinomial:

http://www.ats.ucla.edu/stat/stata/output/stata_mlogit_output.htm

Ejemplos de publicaciones que utilizan la regresión logística binomial:

Aparicio, Ricardo y David H. Corrochano. 2005. El perfil del votante clientelar en México durante las elecciones de 2000. *Estudios Sociológicos* Vol. 23, No. 68 (May - Aug., 2005), pp. 375-396

Cerrutti, Marcela y Rene Zenteno, 2000 “Cambios en el Papel Económico de las Mujeres entre las Parejas Mexicanas”. *Estudios Demográficos y Urbanos*, 15/1

CORTES, F. Y R. RUBALCAVA. (1993), "Algunas determinantes de la inserción laboral en la industria maquiladora de exportación de Matamoros", en: Estudios Sociológicos, Vol. XI, No 31, enero - abril, México, Ed. El Colegio de México.

García, Brígida y Edith Pacheco (2000), "Esposas, hijos e hijas en el mercado de trabajo de la ciudad de México en 1995", Estudios demográficos y Urbanos, El Colegio de México, vol. 15, no. 1, January-April, pp. 35-64

Ejemplos de publicaciones que utilizan la regresión logística multinomial:

López Ramírez, Adriana. 2008. "Migración, remesas y arreglos residenciales de los adultos mayores en México". Estudios Demográficos y Urbanos, Vol. 23, Núm. 3, septiembre-diciembre, pp. 513-541

Andrés Peri e Ignacio Pardo. 2006. "Nueva evidencia sobre la hipótesis de la doble insatisfacción: ¿Cuán lejos estamos de que toda la fecundidad sea deseada?", en III Congreso de la Asociación Latinoamericana de Demografía, Guadalajara (versión en CD-ROM).

Ejemplos de publicaciones que utilizan la regresión logística ordenada:

Solís, Patricio. 2005. "Cambio estructural y movilidad ocupacional en Monterrey, México" Estudios Sociológicos, vol. XXIII, núm. 1, enero-abril, pp. 43-74

Villarreal, Andrés. 2010. "Stratification by Skin Color in Contemporary Mexico". American Sociological Review 75-5.