



PUCP

STATA¹⁷

STATISTICS • VISUALIZATION • DATA MANIPULATION • REPORTING

DEPARTAMENTO DE ECONOMÍA
LABORATORIO DE ECONOMETRÍA: STATA
1ECO31

Sesión 15

Logit Multinomial

Docente: Juan Palomino



Índice

1

¿Cuándo usar Logit Multinomial?

2

Modelo Logit Multinomial

3

Relative Risk Rate y Efectos Marginales

1. ¿Cuándo usar Logit Multinomial

¿Cuándo usar Logit Multinomial?

La variable dependiente “ y ” es una variable categórica no ordenada.

Un individuo puede seleccionar solo una alternativa.

Cuando la empresa escoge entre diferentes alternativas:

- Comerciar sus productos de manera local, regional, internacional
- Modalidad de venta (venta directa a consumidor, venta mediante canal de distribución u otro canal)
- Regimen Tributario (regimen general, regimen especial de impuesto a la renta, regimen mype tributario, otras).

¿Cuándo usar Logit Multinomial?

Cuando el individuo escoge entre diferentes alternativas:

- El tipo de seguro de salud (PACIFICO, RIMAC, MAPFRE, ESSALUD)
- El tipo de sistema de pensión (AFP, ONP, otros, no afiliado).
- El partido político por el cual va a votar en las elecciones (Frente Amplio, PPC, Fuerza Popular)
- Decisiones de vida (estudiar, trabajar, no hacer nada)
- Modo de migrar (avión, tren, bus, carro)

2. Logit Multinomial

Logit Multinomial

- La variable endógena tiene m categorías y se codifican como $j = 1, 2, \dots, m$
- Sea p_1, p_2, \dots, p_m las probabilidades de seleccionar o caer en alguna de estas categorías.
- La probabilidad de que el individuo i elija la alternativa j es:

$$p_{ji} = \frac{\exp(x_i \beta_j)}{1 + \sum_{k=1}^{m-1} \exp(x_i \beta_k)} \quad j = 1, \dots, m-1$$

Categoría base: " m "  $p_{mi} = \frac{1}{1 + \sum_{k=1}^{m-1} \exp(x_i \beta_k)}$

- La decisión por alguna de las alternativas depende de las características x_i

Logit Multinomial

- En el caso $m = 4$, es decir, cuatro categorías:

$$p_{1i} = \frac{\exp(x_i\beta_1)}{1 + \exp(x_i\beta_1) + \exp(x_i\beta_2) + \exp(x_i\beta_3)}$$

$$p_{2i} = \frac{\exp(x_i\beta_2)}{1 + \exp(x_i\beta_1) + \exp(x_i\beta_2) + \exp(x_i\beta_3)}$$

$$p_{3i} = \frac{\exp(x_i\beta_3)}{1 + \exp(x_i\beta_1) + \exp(x_i\beta_2) + \exp(x_i\beta_3)}$$

$$p_{4i} = \frac{1}{1 + \exp(x_i\beta_1) + \exp(x_i\beta_2) + \exp(x_i\beta_3)}$$

La suma de todas las probabilidades debe dar 1.

- Cuando $m = 4$, se estima 3 vectores de parámetros β_1 , β_2 y β_3 .
- Para m alternativas se estimarán $m - 1$ vectores de parámetros.

Forma General

- Se define la variable dependiente $y_{ij} = 1$ si individuo i escoge o cae en la categoría j , e $y_{ij} = 0$ de otro modo.
- Luego, se estiman los $m - 1$ vectores de parámetros, $\beta_1, \beta_2, \dots, \beta_{m-1}$, maximizando la verosimilitud:

$$L = \prod_{i=1}^n p_{i1}^{y_{i1}} p_{i2}^{y_{i2}} \dots p_{im}^{y_{im}}$$

- Tomando logaritmos:

$$\ln L = \sum_{i=1}^n \sum_{j=1}^m y_{ij} \ln p_{ij}$$

- Maximizando esta función se obtienen los estimadores $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_{m-1}$

3. Relative risk rate y Efectos Marginales

Relative risk ratio: Interpretación del coeficiente β

- Tenemos la probabilidad de éxito, $p_i = \frac{\exp(x_i\beta)}{1+\exp(x_i\beta)}$, y de fracaso, $1 - p_i = \frac{1}{1+\exp(x_i\beta)}$
- Dividimos ambas probabilidades, obtenemos

$$\frac{p_i}{1 - p_i} = \exp(x_i\beta)$$

- Tomando logaritmos:

$$\ln\left(\frac{p_i}{1 - p_i}\right) = x_i\beta$$

- Es decir, los β en el logit son el impacto de x sobre el logaritmo de la razón de probabilidades (relative risk ratio)
- Ejemplo: si p_i es la probabilidad de que una empresa comercie internacionalmente y $1 - p_i$ la probabilidad de no comerciar, entonces si $\frac{p_i}{1-p_i} = 5$ significa que es 5 veces más probable de que la empresa comercie internacionalmente a no comerciar.

Relative risk ratio: Interpretación del coeficiente β

El RRR de un coeficiente indica cómo es el riesgo (probabilidad) de que el resultado caiga en la categoría de comparación en relación al riesgo (probabilidad) de que el resultado caiga en la categoría base, cuando cambia con la variable en análisis.

- Un **RRR > 1** indica que la probabilidad de que el resultado caiga en la categoría de comparación en relación con la probabilidad de que el resultado caiga en la categoría base incrementa a medida que aumenta la variable. En otras palabras, el resultado de la comparación es más probable.
- Un **RRR < 1** indica que la probabilidad de que el resultado caiga en la categoría de comparación en relación con la probabilidad de que el resultado caiga en la categoría base disminuye a medida que aumenta la variable. En otras palabras, es más probable que el resultado esté en el grupo de referencia.

Efectos Marginales

- Entrega el efecto de un cambio en la variable sobre la probabilidad absoluta de caer en una de las alternativas.
- A diferencia de los β , su interpretación no es relativa (en comparación con una alternativa base) sino absoluta: el impacto sobre la probabilidad de estar en una categoría.
- Por ello, los efectos marginales no se alteran si se cambia la base.

Referencias

- Greene, W. H., & Hensher, D. A. (2010). Modeling Ordered Choices: A primer. Cambridge University Press.
- Long, J. S., & Freese, J. (2006). Regression Models for Categorical Dependent Variables using Stata. Stata press.
- Multinomial Logistic Regression | Stata annotated Output
<https://stats.oarc.ucla.edu/stata/output/multinomial-logistic-regression/>
- Sarrias, M., & Daziano, R. (2017). Multinomial logit models with continuous and discrete individual heterogeneity in R: the gmn1 package. *Journal of Statistical Software*, 79, 1-46.
<https://www.jstatsoft.org/article/view/v079i02>



PUCP