

Econometría II

Modelos de respuesta binaria

Alejandro Mosiño - *Universidad de Guanajuato*

v. 2026.05.19

Introducción

Introducción (1/4)

Deseamos responder preguntas como las siguientes:

- ¿Por qué algunas personas van a la escuela y otras no?
- ¿Por qué algunas mujeres trabajan y otras no?
- ¿Por qué algunas personas compran casa y otras rentan?
- ¿Por qué algunas personas usan transporte público y otras no?

Introducción (2/4)

Supongamos que estudiamos por qué algunas personas viajan en su automóvil y otras toman el autobús. Pensemos que *no existen otras opciones*. En este caso, podemos definir:

$$y_i = \begin{cases} 1 & \text{Viaja en automóvil} \\ 0 & \text{Viaja en autobús} \end{cases}$$

Sea p_i la probabilidad de que una persona (seleccionada al azar) viaje en automóvil. Esto implica que:

$$\mathbb{P}(y_i = 1) = p_i$$

$$\mathbb{P}(y_i = 0) = 1 - p_i$$

Introducción (3/4)

Nota que la variable aleatoria (Bernoulli) que acabamos de definir tiene una función de probabilidad:

$$f(y_i) = p_i^{y_i} (1 - p_i)^{1-y_i} ; \quad y_i = 0, 1$$

Para esta función tenemos:

$$\mathbb{E}(y_i) = p_i$$

$$\mathbb{V}\text{ar}(y_i) = p_i(1 - p_i)$$

Introducción (4/4)

¿Qué factores pueden afectar el valor de p_i ? Una posibilidad es:

$$x_i = t_i^{bus} - t_i^{auto}$$

donde t_i^{bus} es el tiempo de viaje en autobús y t_i^{auto} es el tiempo de viaje en automóvil. Nota que se espera que p aumente conforme x aumenta, *ceteris paribus*. ¿Cómo podemos modelar esta relación?

El modelo de probabilidad lineal

El modelo de probabilidad lineal (MPL)

Consideremos el siguiente modelo:

$$y_i = \mathbb{E}(y_i) + u_i,$$

donde:

$$\begin{aligned}\mathbb{E}(y_i) &= p \\ &= \beta_1 + \beta_2 x_i\end{aligned}$$

Esto es, nuestro modelo tiene una forma conocida:

$$y_i = \beta_1 + \beta_2 x_i + u_i$$

Limitaciones del MPL (1/4)

Nota que el MPL implica:

$$u_i = \begin{cases} 1 - (\beta_1 + \beta_2 x_i) & \text{si } y_i = 1 \\ -(\beta_1 + \beta_2 x_i) & \text{si } y_i = 0 \end{cases}$$

Por lo tanto:

$$\mathbb{V}\text{ar}(u_i) = (\beta_1 + \beta_2 x_i)(1 - \beta_1 + \beta_2 x_i)$$

¡El error del MPL es heteroscedástico, no es continuo ni normal!

Limitaciones del MPL (2/4)

Ejercicio

Muestra que la varianza del error en el modelo de probabilidad lineal es:

$$\text{Var}(u_i) = (\beta_1 + \beta_2 x_i)(1 - \beta_1 + \beta_2 x_i)$$

Limitaciones del MPL (3/4)

Ahora, nota que en el MPL:

$$\frac{\partial p_i}{\partial x_i} = \beta_2$$

Esto implica, en nuestro ejemplo, que la probabilidad de conducir un automóvil se incrementa (a una tasa constante) conforme x_i se incrementa. El problema es que nada garantiza que:

$$0 \leq p \leq 1$$

Limitaciones del MPL (4/4)

El primer problema del MPL se soluciona estimando los coeficientes por el método de los MCG o por MV. El segundo problema se soluciona suponiendo que:

$$z_i = \beta_1 + \beta_2 x_i ,$$

y que:

$$p = F(z_i)$$

donde $F(z_i)$ es una función sigmoide.

El modelo probit

El modelo probit (1/4)

Para modelar $F(z_i)$ podemos utilizar una **función probit**. Esta se relaciona con una *función de probabilidad normal estándar*, su función de densidad es:

$$\phi(z_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_i^2}$$

El modelo probit es:

Entonces, la función probit es:

$$\begin{aligned}\Phi(z_i) &= \mathbb{P}(Z \leq z_i) \\ &= \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u_i^2} du \\ &= F(z)\end{aligned}$$

Lo que requerimos entonces es:

$$\begin{aligned}p_i &= \mathbb{P}(Z \leq \beta_1 + \beta_2 x_i) \\ &= \Phi(\beta_1 + \beta_2 x_i)\end{aligned}$$

El modelo probit (3/4)

Nota que el efecto marginal de la variable x_i sobre la probabilidad se calcula como:

$$\begin{aligned}\frac{\partial p_i}{\partial x_i} &= \frac{\partial \Phi(\cdot)}{\partial (\cdot)} \\ &= \phi(\beta_1 + \beta_2 x_i) \beta_2\end{aligned}$$

El modelo probit (4/4)

Además:

- Dado que el valor de $\phi(\beta_1 + \beta_2 x_i)$ es una función de densidad, su valor es siempre positivo. Entonces, el valor del efecto marginal depende del signo de β_2 .
- $\phi(\beta_1 + \beta_2 x_i)$ cambia con el valor de x_i . Su valor máximo se alcanza cuando $\beta_1 + \beta_2 x_i = 0$. Además, $p_i = \Phi(0) = 0.5$.
- Si $\beta_1 + \beta_2 x_i$ es muy grande, $\phi(\beta_1 + \beta_2 x_i)$ converge a cero y $p_i = \Phi(\beta_1 + \beta_2 x_i)$ a 1.

Estimación del modelo probit (1/2)

El modelo probit se estima por *máxima verosimilitud*.

Para ver la lógica de la estimación, supongamos que tres individuos son seleccionados al azar: $y_1 = 1$, $y_2 = 1$, $y_3 = 0$, donde $y = 1$ si el individuo viaja en automóvil y $y = 0$ si el individuo viaja en autobús. Además, supongamos que $x_1 = 15$, $x_2 = 6$, $x_3 = 7$. En este caso tenemos:

$$f(y_i) = [\Phi(\beta_1 + \beta_2 x_i)]^{y_i} [1 - \Phi(\beta_1 + \beta_2 x_i)]^{1-y_i} ; \quad y_i = 0, 1, \quad i = 1, 2, 3$$

Y, si los tres individuos son independientes:

$$\begin{aligned} f_{y_1, y_2, y_3} &= f(y_1) f(y_2) f(y_3) \\ &= f(1) f(1) f(0) = \mathbb{P}(y_1 = 1, y_2 = 1, y_3 = 0) \\ &= \Phi(\beta_1 + 15\beta_2) \times \Phi(\beta_1 + 6\beta_2) \times (1 - \Phi(\beta_1 + 7\beta_2)) \\ &= L(\beta_1, \beta_2) \end{aligned}$$

Estimación del modelo probit (2/2)

Ejercicio

Encuentra la función de log-verosimilitud de este ejemplo. Encuentra los estimadores de máxima verosimilitud de β_1 y β_2 . Interpreta tus resultados.

Nota. Para interpretar los resultados de la regresión (es decir, para evaluar los efectos marginales) podemos usar el promedio de las $\mathbf{x}$ en lugar de las $\mathbf{x}_i$.

El modelo probit: Ejemplo (1/3)

En el ejemplo del viaje en automóvil tenemos:

```
##
## Call:
## glm(formula = auto ~ dtime, family = binomial(link = "probit"),
##      data = data)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.06443    0.40068  -0.161  0.87224
## dtime        0.29999    0.10290   2.915  0.00355 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 29.065  on 20  degrees of freedom
## Residual deviance: 12.330  on 19  degrees of freedom
## AIC: 16.33
##
## Number of Fisher Scoring iterations: 7
```

El modelo probit: Ejemplo (2/3)

¿Cuál será la probabilidad de viajar en auto si $dtime = 3$ (30 minutos)?

Primero, calculamos el valor de la función probit, z_i , cuando $dtime = 3$:

$$\begin{aligned} z &= \beta_1 + \beta_2 x \\ &= -0.065 + 0.3dtime \\ &= 0.83 \end{aligned}$$

- Para calcular la probabilidad de viajar en auto, evaluamos la función $F(z_i) = \Phi(z_i)$:

$$\Phi(z_i) = \Phi(0.83) = 0.8$$

- El efecto de un pequeño cambio en $dtime$ cuando $dtime = 3$:

$$\phi(z) \beta_2 = 0.3 \phi(0.83) = 0.085$$

El modelo probit: Ejemplo (3/3)

Nota. Aunque aun es posible calcular el RSS , la R^2 y la R^2 **ajustada**, estos dejan de tener un significado útil en el contexto de los modelos probit. Entonces, es mejor calcular el número de casos predichos correctamente.

En nuestro ejemplo de viaje en auto o en autobús tenemos que este es el **90.5**. ¿Qué implica esto?

El modelo logit

El modelo logit (1/2)

Para modelar $F(\mathbf{z})$ también podemos considerar la **función logística**:

$$p_i = F(z_i) = \frac{1}{1 + e^{-z_i}} = \Lambda(z_i)$$

Nota que:

$$\lim_{z \rightarrow +\infty} F(z) = 1 \quad \text{y} \quad \lim_{z \rightarrow -\infty} F(z) = 0$$

Además:

$$f(z) = \frac{\partial p}{\partial z} = \frac{e^z}{(e^z + 1)^2} = \lambda(z)$$

El modelo logit (2/2)

Lo anterior puede representarse de forma más útil. La probabilidad de que $y = 1$ es:

$$p = \frac{1}{1 + e^{-z}} = \frac{1}{1 + e^{-(\beta_1 + \beta_2 x)}} = \frac{\exp(\beta_1 + \beta_2 x)}{1 + \exp(\beta_1 + \beta_2 x)}$$

Y la probabilidad de que $y = 0$:

$$1 - p = 1 - \frac{\exp(\beta_1 + \beta_2 x)}{1 + \exp(\beta_1 + \beta_2 x)} = \frac{1}{1 + \exp(\beta_1 + \beta_2 x)}$$

Nota. La estimación se hace igual que en el modelo logit, tan solo reemplazamos $\Phi(z)$ por $\Lambda(z)$.

El modelo logit: Ejemplo (1/2)

En el ejemplo del viaje en automóvil tenemos:

```
##
## Call:
## glm(formula = auto ~ dtime, family = binomial(link = "logit"),
##      data = data)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.2376      0.7505  -0.317   0.7516
## dtime          0.5311      0.2064   2.573   0.0101 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 29.065  on 20  degrees of freedom
## Residual deviance: 12.332  on 19  degrees of freedom
## AIC: 16.332
##
## Number of Fisher Scoring iterations: 6
```

Número de resultado correctamente predichos = 90.5.

El modelo logit: Ejemplo (2/2)

Ejercicio

Calcula la probabilidad de viajar en automóvil y el efecto marginal cuando *mtime* = 3. Compara tus resultados con los del modelo probit.

Pruebas de hipótesis (1/3)

En ambos modelos, probit y logit, podemos llevar a cabo hipótesis del tipo:

$$H_0 : \beta_k = c$$

$$H_a : \beta_k \neq c$$

Para hacer esto utilizamos el estadístico t -asintótico o **estadístico de Wald**:

$$t = \frac{\tilde{\beta}_k - c}{s.e.(\tilde{\beta}_k)} \sim t_{(n-k)}$$

Pruebas de hipótesis (2/3)

Normalmente, sin embargo, se prefiere realizar pruebas basadas en el **cociente de verosimilitud. Este es:

$$LR = 2(\ln L_u - \ln L_R),$$

donde L_u es la función de verosimilitud del *modelo no restringido* y L_R es la función de verosimilitud del *modelo restringido*. Si H_0 es correcta, entonces:

$$LR \sim \chi_m^2$$

donde m es el número de restricciones impuestas.

Pruebas de hipotesis (3/3)

Por ejemplo, si en el modelo probit deseamos probar:

$$H_0 : \beta_2 = 0$$

Hipótesis nula: el parámetro de regresión es cero para *dtime*

$$H_a : \beta_2 \neq 0$$

Estadístico de contraste: $F(1,19)=8.50468$, valor p 0.00885597.

Al omitir variables mejoraron 0 de 3 criterios de información.

Y, si deseamos probar:

$$H_0 : \beta_2 = 0$$

$t(19)$: área a la derecha de 2.91628 = 0.00442797)

(valor a dos colas = 0.00885594; complemento = 0.991144)

$$H_a : \beta_2 > 0$$

Modelo probit o logit (1/2)

En la práctica no existen razones reales para preferir uno u otro modelo. Sin embargo, muchos investigadores prefieren el logit debido a que los resultados pueden expresarse en términos del cociente de probabilidades (cociente de momios, odds ratio):

$$\frac{p}{1-p} = e^{z_i}$$

Modelo probit o logit (2/2)

En nuestro ejemplo:

Odds-ratios for auto:

Variable	Odds-ratio	95.0% conf. interval
dtime	1.7008	[1.135, 2.549]

Si *dtime* aumenta en 10 minutos, la probabilidad de viajar en auto en **70.08** más grande que la de viajar en autobús.

Práctica

Práctica

Utiliza los datos: [11-microeconometria_datos-smoke](#) para calcular la probabilidad de convertirse en fumador. La muestra contiene 1196 hombres. Calcula probabilidades y los efectos marginales utilizando modelos de probabilidad lineal, logit y probit. Los datos que tenemos son:

smoker : Variable dummy = 1 si la persona fuma y cero si no.

age : Edad (en años).

educ : Escolaridad (en años).

income : Ingreso familiar.

pcigs79 : Precio de los cigarrillos en Estados Unidos(1979).

Nota. La base de datos también contiene unas variables interactivas. Puedes incluirlas en tu modelo.