

ECONOMETRIA II – CURSO 2016

PRÁCTICO 5
Modelos Binarios

EJERCICIO 1

A partir de una encuesta de marketing sobre consumo de bebidas dietéticas, se observó que de 1000 personas encuestadas (570 mujeres y 430 hombres), el 63% prefería esta categoría de bebidas a las tradicionales. Esta preferencia resultó ser sensiblemente superior entre las consumidoras mujeres, de las 570 encuestadas, 450 respondieron afirmativamente. Por el contrario, entre los 430 encuestados de sexo masculino, sólo 180 manifestaron preferirlas.

Para analizar este comportamiento particular se especificó el modelo econométrico siguiente:

$$P(Y = 1 | MUJER) = F(MUJER, a_0, a_1) = \frac{1}{1 + e^{-(a_0 + a_1 MUJER)}}$$

donde: $Y = 1$ si el individuo prefiere las bebidas dietéticas, y $Y = 0$ en otro caso.

$MUJER = 1$ si la persona es mujer, 0 si es varón.

Se estimaron por el método de máxima verosimilitud los parámetros del modelo anterior, y después de un modelo restringido donde $a_1 = 0$, obteniendo los resultados siguientes:

[1] Modelo sin restricción

Coef.	Variable	Estim.	Err.Est.
a_0	Constante	-0.328	0.031
a_1	MUJER	1.650	0.448

$$\ln[L(a)] = -585.68$$

[2] Modelo con restricción $a_1 = 0$

Coef.	Variable	Estim.	Err.Est.
b_0	Constante	0.5322	0.207

Se pide:

1. Escribir la verosimilitud del modelo Logit en este caso.
2. Analizar los resultados de la estimación en [1] explicitando el valor del estadístico para la prueba de significación de los coeficientes, la significación estadística y la interpretación de los estimadores.
3. Calcular $\ln[L(b_0)]$ para el Modelo [2] y contrastar la hipótesis nula $a_1 = 0$
4. Mostrar que el modelo binario es intrínsecamente heteroscedástico. Explicitar los problemas de estimación y de formulación que se presentan cuando se estima un modelo de probabilidad lineal.
5. ¿Qué sucedería si en lugar de 1000 observaciones contáramos con 100 igualmente distribuidas entre hombre/mujeres y consumidores/no consumidores?

EJERCICIO 2

Ejercicio basado en *Sección 7.5 y Ejemplos 8.8 y 17.1 de Wooldridge (2004) y Ejemplo 15.2 de Wooldridge (2002), utilizando los datos de Mroz (1987)*

Para estimar la participación en la fuerza de trabajo de las mujeres casadas se propone el siguiente modelo:

$$\text{inlf}_i = \beta_0 + \beta_1 \text{nwifeinc}_i + \beta_2 \text{educ}_i + \beta_3 \text{exper}_i + \beta_4 \text{exper2}_i + \beta_5 \text{age}_i + \beta_6 \text{kidslt6}_i + \beta_7 \text{kidsge6}_i + u_i$$

donde:

<i>inlf</i>	=1 si participa de la fuerza de trabajo en 1975
<i>nwifeinc</i>	Ingreso del resto de los miembros del hogar (en miles de dólares)
<i>educ</i>	Educación en años
<i>exper</i>	Experiencia laboral en años
<i>exper2</i>	Experiencia laboral en años, al cuadrado
<i>age</i>	Edad en años
<i>kidslt6</i>	Cantidad de hijos menores de 6 años de edad
<i>kidsge6</i>	Cantidad de hijos entre 6 y 18 años de edad

Se estima un MPL y modelos Probit y Logit.

Variable dependiente inlf ("In Labor Force")				
	MPL (1)	LOGIT	PROBIT	
nwifeinc	-0.0034	-0.0213	-0.012	
s.d.^	0.0015	0.0084	0.0048	
t stat.	-2.2	-2.5	-2.5	
Educ	0.038	0.2212	0.1309	
s.d.^	0.0073	0.0434	0.0252	
t stat.	5.2	5.1	5.2	
Exper	0.0395	0.2059	0.1233	
s.d.^	0.0058	0.0321	0.0187	
t stat.	6.8	6.4	6.6	
expersq	-0.0006	-0.0032	-0.0019	
s.d.^	0.0002	0.001	0.006	
t stat.	-3.1	-3.1	-3.1	
Age	-0.0161	-0.088	-0.0529	
s.d.^	0.0024	0.0146	0.0085	
t stat.	-6.7	-6.0	-6.2	
kidslt6	-0.2618	-1.4434	-0.8683	
s.d.^	0.0317	0.2036	0.1185	
t stat.	-8.2	-7.1	-7.3	
kidsge6	0.013	0.0601	0.036	
s.d.^	0.0135	0.0748	0.0435	
t stat.	1.0	0.8	0.8	
_cons	0.5855	0.4255	0.2701	
s.d.^	0.1523	0.8604	0.5086	
t stat.	3.9	0.5	0.5	
Numero de Observaciones	753	753	753	
Porcentaje correctamente predicho	73.4	73.6	73	
Log de la verosimilitud	----	-401.77	-401.30	
Pseudo R-cuadrado	0.264	0.220	0.221	

(1) Los errores estandar estan corregidos por heterocedasticidad

Se pide:

1. ¿Cuáles son las variables que determinan la decisión de las mujeres de participar en el mercado de trabajo según el Modelo de Probabilidad Lineal? ¿Y según los modelos Probit y Logit? ¿Depende esta conclusión de los datos particulares de este problema?
2. ¿Cómo explica la diferencia entre las estimaciones de los coeficientes del MPL y del Modelo Probit? ¿Qué consecuencias tiene esta diferencia en la estimación de los efectos parciales?
3. ¿Cómo explica la diferencia entre las estimaciones de los coeficientes del Modelo Logit y del Modelo Probit? ¿Qué consecuencia tiene esta diferencia en la estimación de los efectos parciales?
4. ¿Cuál de los tres modelos elegiría para estimar el efecto que tiene un año de educación adicional en la probabilidad de que una mujer con 12 años de educación participe en el mercado de trabajo?
5. ¿Modificaría su respuesta en 4 si le dijeran que estimara el efecto para una mujer con 5 años de educación?
6. ¿Cómo cambia la probabilidad de que una mujer casada trabaje si pasa de tener 1 hijo de 6 años o menos a tener 3 hijos en dicho tramo de edad? Considere el caso de una mujer con 32 años de edad, 12 años de educación, 8 años de experiencia laboral, que no tiene hijos mayores de 6 años y que vive en un hogar en el cual el resto de los miembros tiene un ingreso de 20 (en logaritmos).

EJERCICIO 3

(Basado en el Problema 15.3 de Wooldridge, J. M. (2010), *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Estados Unidos, 2ª Edición, página 636.)

Sean z_1 y z_2 variables continuas y d_1 una variable ficticia.

Considere el siguiente modelo Probit

$$P(y = 1 | z_1, z_2) = \Phi(\alpha + \delta_1 z_1 + \gamma_1 z_2 + \gamma_2 z_2^2)$$

Se pide 1:

1. Escriba los efectos parciales de z_1 y z_2 .
2. Considere ahora una especificación diferente del modelo, incluyendo la variable ficticia d_1

$$P(y = 1 | z_1, z_2, d_1) = \Phi(\alpha + \delta_1 z_1 + \gamma_1 z_2 + \gamma_2 d_1 + \gamma_3 z_2 d_1)$$

Se pide 2:

1. Escriba el efecto parcial de z_2 .
2. Escriba el efecto parcial de d_1 . ¿Cómo lo estimaría?

EJERCICIO 4

Problema 15.7 en Wooldridge, J. M. (2010), *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Estados Unidos, 2ª Edición, página 638.

Utilice los datos del archive CRIME.DTA para esta pregunta.

- a. Defina una variable binaria, llamada *arr86*, igual a la unidad si un hombre fue arrestado al menos una vez durante 1986, y cero en otro caso. Estime un MPL que relacione *arr86* con *pcnv*, *avgsen*, *tottime*, *ptime86*, *inc86*, *black*, *hispan* y *born60*. Reporte los errores estándar habituales y los robustos a la heteroscedasticidad. ¿Cuál es el efecto estimado sobre la probabilidad de arresto si *pcnv* pasa de 0.25 a 0.75?
- b. Pruebe la significación conjunta de *avgsen* y *tottime*, usando una prueba no robusta y una prueba robusta.
- c. Ahora estime el modelo mediante un Probit. En la media muestral de *avgsen*, *tottime*, *inc86*, y *ptime86*, y con *black*=1, *hispan*=0, y *born60*=1, ¿Cuál es el efecto estimado sobre la probabilidad de arresto si *pcnv* pasa de 0.25 a 0.75? Compare este resultado con su respuesta a la parte a.
- d. Para el modelo Probit estimado en la parte c, obtenga el porcentaje de predicciones correctas. ¿Cuál es el porcentaje predicho correctamente cuando *narr86*=0? ¿Cuándo *narr86*=1? ¿Cómo interpreta estos resultados?
- e. En el modelo Probit, agregue los términos $pcnv^2$, $ptime86^2$, y $inc86^2$ al modelo. ¿Son individual o conjuntamente significativos? Describa la relación estimada entre la probabilidad de arresto y *pcnv*. En particular, ¿en qué punto sucede que la probabilidad de ser convicto tiene un efecto negativo sobre la probabilidad de arresto?

EJERCICIO 5

Problema 17.2 en Wooldridge, J. M. (2006), *Introducción a la Econometría: un enfoque moderno*, Thomson Learning, México, 2ª Edición, página 664.

EJERCICIO 6

Ejercicio 17.9 en Wooldridge, J. M. (2006), *Introducción a la Econometría: un enfoque moderno*, Thomson Learning, México, 2ª Edición, página 666. Utilizar datos del archivo LOANAPP.DTA.