MICROECONOMETRÍA: MODELOS DE RESPUESTA BINARIA

Norma Edith Alamilla López*, Sigfredo Arauco Camargo*.

Alamilla-López N.E., Arauco-Camargo S.
Microeconometria: Modelos de respuesta binaria.
Hitos de Ciencias Económico Administrativas
2009;15 (42): 83-88.

RESUMEN

Un modelo de respuesta binaria es una regresión en la cual la variable dependiente es una variable aleatoria que toma valores de cero y uno que puede estar explicada por otras variables independientes incluyendo una variable de perturbación aleatoria que recoge las desviaciones que los agentes tienen respecto a lo que sería el comportamiento del agente medio. Estos modelos de decisión se aplican en la gestión empresarial, por ejemplo: Cuando un empresario desea lanzar un nuevo producto al mercado si la utilidad que le proporciona este hecho supera a la de no hacerlo, o cuando el director de un banco decidirá conceder un préstamo si la utilidad de concederlo supera la de no hacerlo. En base a estos ejemplos, el planteamiento teórico de las preferencias de los agentes económicos permite ordenar las diferentes alternativas en función de su atractivo para el individuo decisor en los que se fundamentan los modelos de elección discreta.

Palabras Clave: Modelos logit. Modelos probit. Máxima verosimilitud. Liquidez.

Alamilla-López N.E., Arauco-Camargo S.
Microeconometrics: Binary response models.
Hitos de Ciencias Económico Administrativas
2009;15 (42): 83-88.

ABSTRACT

A binary response model is a regression in which the dependent variable is a random variable that takes values of zero and one and can be explained by other independent variables including a variable of random disturbance that gathers the diversions that the agents have with regard to what would be the behavior of the average agent. These models of decision are applied in the business management, for example: when a businessman wants to launch a new product to the market if the utility he receives is more than not doing it, or when the director of a bank decides to grant a loan, the utility of granting it is more than not granting it. On the basis of these examples, the theorizing of the preferences of the economic agents allows it to arrange the different alternatives depending on their attractiveness for the individual decisor which the models of discreet choice are based on.

Key words: Logit models. Probit models. Maximum verisimilitude. Liquidity.

DIRECCIÓN PARA RECIBIR CORRESPONDENCIA: Correo electrónico: norma_alamilla@hotmail.com

Los fines u objetivos de la empresa pueden ser múltiples, mantener competitividad, absorber nuevos segmentos de mercados, modernizarse, etc. Para lograr lo anterior se debe de interpretar de manera operativa a la empresa mediante la consecución de valores añadidos en sentido financiero que se relacionan con la rentabilidad-seguridad y la solvencia-estabilidad que se derivan de las razones financieras o comparación por cociente de valores con significado para la empresa que sirven para la toma de decisiones.

Tradicionalmente, la econometría se enfrentaba a los problemas de decisión utilizando los modelos agregados, de forma que haciendo uso del modelo general se resolvía la conducta esperada de los agentes económicos. Pero cuando se quiere captar los procesos de decisión individual y las relaciones causales propias de la economía o empresa en los procesos de decisión es vital medirlo en un contexto individual.

* Profesores investigadores de la Universidad Tecnológica de la Mixteca.

Fecha de recibido: 5 de septiembre 2008. **Fecha de aceptación:** 1 de diciembre 2008.

En dichos estudios de microeconometría se necesita un marco de información con datos individuales, es decir datos desagregados para cada uno de los agentes económicos, lo cual supone mucho trabajo en la generación de la información.

En la primera parte se desarrollan los modelos de elección binaria que pueden ser de tipo dicotómica o polinómicas. En la segunda, parte se desarrollan de manera particular los modelos Logit y Probit; que típicamente son estimados por el método de máxima verosimilitud, debido a que este estimador tiene buenas propiedades. En particular, es asintóticamente eficiente, es decir, es un estimador mas preciso. En el apartado tres se desarrolla un ejemplo de su aplicación del grupo CEMEX permitiendo con ello elegir aquel modelo que tiene mayor consistencia estadística, y finalmente realizamos las conclusiones pertinentes.

Modelos de elección binaria.

En los modelos de elección binaria se supone que los individuos se enfrentan con una elección entre dos alternativas y que la elección depende de características identificables. En esta situación, la variable endógena puede tomar dos valores: $Y_t = \{0,1\}$ y se pretende explicar la elección hecha por el decisor como función de unas variables que le caracterizan y que se denota por x_i un vector de dimensión *k.*

El propósito de un modelo de elección cualitativa es determinar la probabilidad de que un individuo con un conjunto determinado de atributos hará una elección en lugar de la alternativa. De manera más general, lo que se pretende es encontrar una relación entre un conjunto de atributos que describen a un individuo y la probabilidad de que el individuo hará la elección determinada.

La población en general, y los empresarios, en particular, se encuentran en situaciones en que deben de elegir o decidir entre posibles alternativas. Suponiendo que esas alternativas fueran solo dos opciones, el proceso de modelización adquiere un carácter especial denominándose modelos de respuesta dicotómica o binaria si este fuera el caso, hay incluso modelos multinomiales de más de dos opciones. Algunos ejemplos que se pueden plantear en este tipo de modelización son los siguientes:

- Un trabajador económicamente activo puede estar en situación de paro o trabajando.
- Un trabajador puede plantearse afiliarse o no a un sindicato.
- Un agente económico decide suscribir o no una póliza de seguros.
- Una entidad financiera o banco se encuentra ante la decisión de conceder o no un crédito (familia, empresa, corporación, etc).
- El área de ventas de una empresa puede estudiar la probabilidad de si un crédito concedido a un cliente será devuelto o no en la fecha de vencimiento.

El planteamiento de este modelo se fundamenta en la siguiente ecuación:

$$Y_t = x_i^t \beta + \varepsilon$$

Dada las opciones:

$$Y_i = \begin{cases} 1 \\ 0 \end{cases}$$

La especificación de la decisión se establecería a través de la ecuación:

$$\Pr ob(Y = 1) = x_i^t \beta$$

Modelos Probit y Logit.

Dadas las dificultades asociadas con el modelo lineal de probabilidad, es natural transformar el modelo original de tal forma que las predicciones caigan en el intervalo [0, 1]. Es decir, ára asegura que P caiga entre 0 y 1, se requiere una función monótona positiva que mapee el predictor lineal $\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 = x^t \beta$ al intervalo [0,1]. Entonces debería adoptarse un modelo bajo el cual los valores de P_i estén restringidos al intervalo [0,1]. Una forma muy convincente de restringir la forma funcional es la siguiente:

$$P_i(\eta_i) = F(x_i^t \beta)$$

En donde $F(\cdot)$ es una función de distribución acumulada, (FDA). La cual es la función diferenciable monótona creciente con dominio R y rango [0,1].

El modelo no lineal sería el siguiente:

$$Y_i = F(x_i^t \beta) + \varepsilon_i$$

con:

$$\varepsilon_i = E(Y_i \mid x_i) - F(x_i^t \beta)$$

Algunas características de la función $F(x_i^t \beta)$

1. Obviamente se trata de una función no lineal, pero una muy particular, en el sentido de que las variables exógenas afectan la variable endógena a través de un índice lineal $x_i^t \beta$, que luego es transformado por la función $F(\cdot)$ de manera tal que los valores de la misma sean consistentes con los de una probabilidad.

¿Cómo elegir la función *F(·)*?
La función de distribución acumulada de cualquier variable aleatoria continua tiene la propiedad de *F(·)*.

Primeramente, si se elige a *F(·)* como la distribución uniforme acumulada entonces obtenemos la construcción del modelo de probabilidad lineal.

Aunque son posibles varias alternativas de la FDA, solo se considerarán dos: La normal y la logística.

El modelo de probabilidad probit se asocia con la función de distribución normal acumulada.

$$\pi_i = \Phi(x_i^t \beta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_i^t \beta} e^{-\frac{z^2}{2}} dZ$$

Donde Z es la variable normal estándar, es decir $Z \sim N(0,1)$.

Además se tiene que:

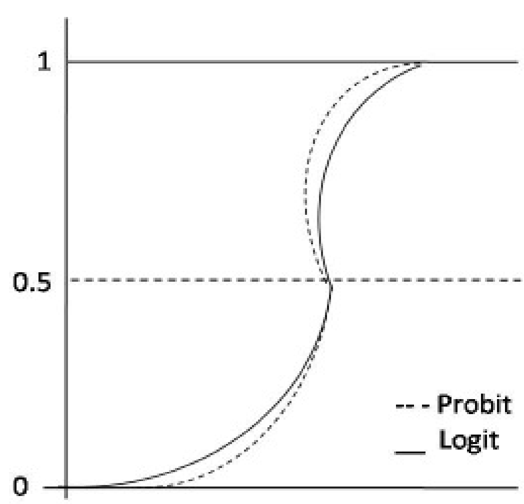
$$\pi_i = \Phi(x_i^t \beta) \Rightarrow \Phi^{-1}(\pi) = x_i^t \beta.$$

Usando la distribución logística $\Lambda(\cdot)$ se produce el modelo logit lineal.

$$\pi_i = \Lambda(x_i^t \beta) = \frac{1}{1 + e^{-x_i^t \beta}} = \frac{e^{x_i^t \beta}}{1 + e^{x_i^t \beta}}$$

La gráfica 1 muestra las gráficas de las distribuciones normal y logística.

GRÁFICA 1
FUNCIÓN DE DISTRIBUCIÓN ACUMULADA NORMAL Y LOGÍSTICA



Las diferencias básicas entre estas dos funciones con forma de “S” invertida, residen en el comportamiento de las colas, esto es, para valores próximos a 0 o a 1, (Gráfica 1).

Dada la similitud existente entre las curvas de la normal acumulada y de la acumulada logística, los resultados estimados por ambos modelos no difieren mucho entre sí, ya que puede apreciarse en la figura 1 que discrepan, únicamente en la rapidez con que las curvas se aproximan a los valores extremos y por lo tanto la función logística es mas achatada que la normal, al alcanzar ésta última más rápidamente los valores extremos, 0 y 1.

A pesar de su similitud, existen dos razones prácticas que aventajan al modelo Logit:

1. Simplicidad: la ecuación de la FDA logística es muy simple, mientras que la FDA normal involucra una integral que no es fácil de evaluar.
2. Interpretabilidad: Una interpretación más sencilla del parámetro estimado es la que se obtiene a través de la linealización del modelo

$$\ln\left(\frac{\pi_i}{1-\pi_i}\right) = \ln\left(\frac{\frac{1}{1+e^{-x_i^t\beta}}}{\frac{1}{1-\frac{1}{1+e^{-x_i^t\beta}}}}\right) = \ln\left(\frac{\frac{e^{x_i^t\beta}}{1+e^{x_i^t\beta}}}{\frac{1}{1+e^{x_i^t\beta}}}\right) = \ln(e^{x_i^t\beta}) = x_i^t\beta$$

Al cociente entre la probabilidad de que ocurra un hecho frente a la probabilidad de que no suceda, se le denomina “odds ratio”. Su interpretación es la preferencia de la opción 1 frente a la opción 0, es decir, el número de veces que es más probable que ocurra un fenómeno frente a que no ocurra.

Odds ratio=. $\ln\left(\frac{\pi_i}{1-\pi_i}\right)$

Dada una variable aleatoria, caracterizada por unos parámetros, y dada una muestra poblacional, se consideran estimadores Máximo-Verosímiles de los parámetros de una población determinada, aquellos valores de los parámetros que generarían con mayor probabilidad la muestra observada. Es decir, los estimadores Máximo-Verosímiles son aquellos valores para los cuales la función de densidad conjunta (o función de verosimilitud) alcanza un máximo.

Estimación de los parámetros en el modelo Probit.
En el caso del modelo Probit, la función de verosimilitud es:

$$L = \prod_{i=1}^{n} \left[\Phi(x_i^t\beta)\right]^{Y_i} \left[1-\Phi(x_i^t\beta)\right]^{1-Y_i}$$

aquí puede apreciarse que, para cada individuo , el termino correspondiente en la función de verosimilitud es, simplemente, $\Phi(x_i^t\beta)$ ó $1-\Phi(x_i^t\beta)$,dependiendo de que $Y_i = 1$ ó $Y_i = 0$. Por tanto, considerando el logaritmo de la verosimilitud se tiene que:

$$\ln L = \sum_{i=1}^{n} Y_i \ln\left[\Phi(x_i^t\beta)\right] + \sum (1-Y_i)\ln\left[1-\Phi(x_i^t\beta)\right]$$

y, calculando derivadas con respecto al vector , se tiene que:

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^{n} Y_i \frac{f(x_i^t\beta)}{\Phi(x_i^t\beta)} x_i + \sum_{i=1}^{n} (1-Y_i)\frac{-f(x_i^t\beta)}{1-\Phi(x_i^t\beta)} = 0$$

donde: $f(x_i^t\beta) = \Phi'(x_i^t\beta)$

o equivalente

$$\sum_{i=1}^{n} \frac{Y_i - \Phi(x_i^t\beta)}{\Phi(x_i^t\beta)\left[1-\Phi(x_i^t\beta)\right]} f(x_i^t\beta) x_i = 0$$

El cual trata de un sistema de k- ecuaciones no lineales.

Estimación de los parámetros en el modelo Logit.
Suponiendo que las observaciones son independientes, la función de la densidad conjunta de las variables dicotómicas Y_i queda como:

$$L = \prod_{Y_i=1} F(x_i^t\beta) \prod_{Y_i=0}\left[1-F(x_i^t\beta)\right] = \frac{e^{\left(\sum_{i=1}^{n} Y_i(x_i^t\beta)\right)}}{\prod_{i=1}^{n}\left(1+e^{x_i^t\beta}\right)}$$

o, lo que es lo mismo:

$$\ln L = \sum_{i=1}^{n} Y_i (x_i^t\beta) - \sum_{i=1}^{n} Ln(1+e^{x_i^t\beta}) = \sum_{i=1}^{n}\left[(Y_i x_i^t)\beta - Ln(1+e^{x_i^t\beta})\right]$$

y detonado por $Z^t = \sum_{i=1}^{n} Y_i x_i^t$, un vector fila 1x*k, se tiene:*

$$\ln L = Z^t\beta - \sum_{i=1}^{n} Ln(1+e^{x_i^t\beta})$$

y,

$$\frac{\partial \ln L}{\partial \beta} = Z - \frac{\sum_{i=1}^{n} x_i e^{x_i^t\beta}}{1+e^{x_i^t\beta}} = 0$$

El cual se trata de un sistema de *k*-ecuaciones no lineales por lo que es necesario aplicar un método iterativo o algoritmo de optimización que permita la convergencia de los estimadores.

Aplicación al caso del grupo CEMEX.
El problema de la decisión entre dos o mas alternativas implica el planteamiento de un modelo que considera una variable dependiente que no es cuantitativa, es decir, no tiene un valor concreto, sino que es cualitativa, y se codifica mediante dígitos o categorías, lo que implica elaborar el MLP y modelos binarios.

Aplicando el modelo Logit.
Al ajustar el modelo Logit a la variable exógena ventas, usando E-Views, las estimaciones que proporciona son:

$$\pi_i = \Lambda(x_i^t\beta) = \frac{1}{1+e^{-(x_i^t\beta)}} = \frac{e^{(x_i^t\beta)}}{1+e^{-(x_i^t\beta)}}$$

Dependent Variable: LIQ
Method: ML - Binary Logit (Quadratic hill climbing)
Sample: 1999:1 2007:1
Included observations: 33
Convergence achieved after 11 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-4.604140	1.404722	-3.277616	0.0010
VTAS	1.78E-06	6.57E-07	2.716337	0.0066

Mean dependent var	0.363636	S.D. dependent var	0.488504
S.E. of regression	0.349894	Akaike info criterion	0.875820
Sum squared resid	3.785198	Schwarz criterion	0.966517
Log likelihood	-12.45102	Hannan-Quinn criter.	0.906336
Restr. log likelihood	-21.63090	Avg. log likelihood	-0.377304
LR statistic (1 df)	18.35975	McFadden R-squared	0.424387
Probability(LR stat)	1.83E-05		

Todos los coeficientes del modelo resultan significativos al 95% en función del estadístico de Wald (columna "z-Statistic" en la tabla anterior), como puede comprobarse a través de sus p-valores asociados (columna "Prob"). El modelo ajustado resulta:

$$\log it(\hat{\pi}_i) = \frac{e^{(-4.604140+1.7E-06*vtas)}}{1+e^{(-4.604140+1.7E-06*vtas)}}$$

Si atendemos a los contrastes basados en los criterios de información de Akaike, Schwarz y de Hannan-Quinn establecen que mientras más bajos sean sus valores mejor será el modelo. En la tabla anterior se muestra que estos criterios tienen un valor muy pequeños 0.875820, 0.966517, y 0.906336, respectivamente. A tenor de ellos, podemos decir que el modelo cumple satisfactoriamente este requisito, y que por lo tanto es un buen modelo.

Ahora, si nos basamos en el hecho, de que "predecir correctamente" equivale a que los ceros sean estimados con una P<0.5, y los unos con una P>0.5, se puede calcular el porcentaje de acierto, lo cual nos da como resultado la siguiente tabla:

		Predichos	
		No	Si
Observados	No	95	5
	Si	33.33	66.67

Es decir, de alguna manera esta tabla nos esta indicando que son "bien" predichos un 95% de las variables endógenas $Y_i = 1$ y un 66.67% de las variables endógenas tales que $Y_i = 0$.

Aplicando el modelo Probit.

Al ajustar el modelo Probit a la variable exógena ventas, usando E-Views, todos los coeficientes del modelo resultan significativos al 95% en función del estadístico de Wald, como puede comprobarse a través de sus *p*-valores asociados. Los resultados se muestran en la siguiente tabla:

Dependent Variable: LIQ
Method: ML - Binary Probit (Quadratic hill climbing)
Date: 09/26/07 Time: 10:50
Sample: 1999:1 2007:1
Included observations: 33
Convergence achieved after 11 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-2.709290	0.745706	-3.633191	0.0003
VTAS	1.06E-06	3.55E-07	2.980060	0.0029

Mean dependent var	0.363636	S.D. dependent var	0.488504
S.E. of regression	0.350820	Akaike info criterion	0.873079
Sum squared resid	3.815313	Schwarz criterion	0.963776
Log likelihood	-12.40580	Hannan-Quinn criter.	0.903595
Restr. log likelihood	-21.63090	Avg. log likelihood	-0.375933
LR statistic (1 df)	18.45020	McFadden R-squared	0.426478
Probability(LR stat)	1.74E-05		

$$\pi_i = \Phi(x_i^t \beta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x_i^t \beta)} e^{-\frac{z^2}{2}} dZ.$$

El modelo ajustado resulta:

$$\Phi^{-1}(\hat{\pi}_i) = -2.709290 + 1.60E - 0.6 * vtas$$

Esto aplica que:

$$\hat{\pi}_i = \Phi(-2.709290 + 1.60E - 0.6 * vtas)$$

Al igual que el modelo Logit, si atendemos a los contrastes basados en los criterios de información de Akaike, Schwarz y de Hannan-Quinn que establecen que mientras más bajos sean sus valores mejor será el modelo, se tiene que los valores que se obtienen son 0.873079, 0.963776, y 0.903595, respectivamente. A tenor de ellos, podemos decir que el modelo cumple satisfactoriamente este requisito, y que lo tanto también es un buen modelo.

Ahora si nos basamos en el hecho, de que "predecir correctamente" equivale a que los ceros sean estimados

		Predichos	
		No	Si
Observados	No	95	5
	Si	33.33	66.67

con una P<0.5, y los unos con una P>0.5, se puede calcular el porcentaje de acierto, lo cual nos da como resultado la siguiente tabla:
Es decir, de alguna manera esta tabla me esta indicando que son «bien» predichos un 95% de la variables endógenas y un 66.67% de las variables endógenas tales que

El cual da exactamente los mismos resultados que en el caso del modelo Logit. Dado que los dos modelos tienen resultados similares ante tratamientos diferentes es importante realizar la validación y contraste de hipótesis con la finalidad de elegir el mejor modelo entre Logit y Probit. El criterio general para la elección entre dos modelos es el siguiente:

La función de verosimilitud (Log likehood). Se refiere a aquel modelo que presente un valor de la función verosimilitud mayor, siendo el modelo Logit que cumple esta condición.

Akaike (1973), sirve para comparar la bondad del ajuste entre dos modelos. Según el criterio es preferible aquel modelo que presente un valor Akaike menor, el modelo Logit tiene un valor de .8758 que es mayor que .8730.

Schwarz, es una propuesta alternativa a la prueba anterior, el criterio es que tenga el menor valor, en el modelo Logia tiene un valor de .9637 a diferencia del modelo Probit que tiene .9637.

Por último, según Hannan-Quinn (1979) es preferible aquel modelo que presente un valor del estadístico menor. El modelo Logit tiene un valor de .9063 mientras que el modelo Probit es de .9035.

La medida de bondad de ajuste, dado que las pruebas tradicionales de bondad de ajuste, tales como el coeficiente de determinación R^2, no son validas en los modelos en los que las variables endógenas toma exclusivamente los valores de uno o cero, se utiliza el r^2 propuesto por McFadden (1974), este estadístico no tiene una interpretación tan directa como el r^2 del modelo de regresión lineal. Se considera el mayor valor del estadístico, por lo tanto, el modelo Logit tiene un valor de .4243 mientras que el modelo Probit es de .4264.

La evaluación anterior determina que el modelo Probit es el que mejor explica estadísticamente la relación entre la liquidez y las ventas.

CONCLUSIONES.

Este estudio examina la aplicación de modelos dicotómicos como los modelos Logit y Probit a partir de las limitaciones del modelo lineal probabilístico en términos teóricos, para posteriormente convalidar la exposición mediante la aplicación de un modelo sencillo de decisión empresarial para el caso de la empresa CEMEX, poniendo a prueba el modelo de decisión en los dos métodos de modelación.

Se exponen los modelos Logit y Probit. En este caso se revisa su especificación, estimación y la evaluación de la bondad del modelo confirmando teóricamente y empíricamente la consistencia de los modelos para estudios de decisión de tipo empresarial.

De acuerdo con el modelo especificado para la empresa CEMEX, su decisión se explica entre otros factores, las ventas, el modelo va a explicar la existencia de liquidez en la empresa en términos dicotómicos.

Los modelos de decisión binaria, aquí estudiados tienen buen ajuste y definición. En ambos casos los resultados son estadísticamente significativos, por tanto, se acepta cualquiera de los dos modelos, dado que existe causalidad en el contexto empresarial ya que se ha representado la probabilidad de tener liquidez para distintos niveles de renta de un potencial incremento o decremento de ventas.

REFERENCIAS

Akaike, H (1973). *Information theory and extensions of the maximum likehood principles*. En B.N. Petrov y F Csaki (Eds). Second International Symposiumon information theory. Akademia Kiada: Budapest.

EVIEWS v.3.1. User's Guide (1997). Quantitative Micro Software. California.

Finney, D.J. (1971). *Probit analysis*. (3ra. ed.). Cambrige University Press.

Gujarati, D. N. (1995). *Basic econometrics*. (3ra. ed.). México: McGraw-Hill.

Hannan, E.J., and B.G. Quinn. (1979). *The determination of the order of an autoregression*. Journal of the Royal Statistical Society, Series B 41:190–195.

Intriligator, M.D. (1990). *Modelos econométricos, técnicas y aplicaciones*. México: Fondo de Cultura Económica S.A.

Lawrence J. Gitman (2003). *Administración financiera*. México: Pearson.

McFadden, D.L. (1974). Conditional logit analysis of qualitative choice behavior. En Zarembka, P. (Ed): Frontiers in Econometrics, Academic Press, New York. pp. 105-142.

Novales, A. (1993). *Econometría*. (2da. ed.). US: McGraw-Hill.

Pérez, Alfredo. (2000). *Los Estados financieros, su análisis e interpretación*. México: ECAFSA.

Pulido, A. (1989). *Predicción económica y empresarial*. Madrid, España: Pirámide.

Schwarz, G. (1978). *Estimating the dimension of a mode*. *Annals of statistics* 6(2):461-464.