

Módulo de estudio sobre Modelos Probit y Logit

María E. Enchautegui*

I. Prefacio

Los modelos probit y logit se han proliferado en la literatura económica desde la década de 1980 en adelante. Los economistas, particularmente los interesados en microeconomía, se han interesado enormemente en este tipo de modelo como marco para estudiar las decisiones que hacen los agentes económicos. El interés en modelos de decisión se debe en gran parte a la mayor disponibilidad de datos micro-económicos en las últimas décadas y a la facilidad de su análisis gracias a la computadora personal. Se entiende que estudiar las decisiones de los agentes económicos es importante de por sí, que estas decisiones son importantes para explicar la conducta de variables agregadas y que son importantes para estudiar su efecto en otras variables de interés.

A pesar de su creciente presencia en la literatura el tratamiento de libros de texto de econometría sobre este tema es parco. Los textos proveen poca discusión y detalle, pocas aplicaciones y las presentaciones del tema no son muy pedagógicas. Por ejemplo, los textos de econometría en su mayoría no discuten la interpretación de coeficientes. Textos especializados en el tema caen al otro extremo, haciendo su lectura difícil para la que solo quiere relacionarse casualmente con estos modelos y los que los necesitan desde un punto aplicado.

Este documento agrupa material sobre los modelos probit y logit que se encuentran dispersos en distintos textos. Además presenta este material en forma concisa para que sirva de introducción al estudio de estos modelos. El material se presenta en forma accesible para personas interesadas en su uso aplicado. Para destacar la aplicación de modelos probit y logit, la sección X se dedica a discutir ejemplos. Además se incluye una lista básica de referencias -la cual no trata de ser exhaustiva- para aquellos lectores interesados en profundizar más sobre en los modelos de decisión binaria.

II. Introducción

En los modelos probit y logit sencillos la variable dependiente toma valores de 0 ó de 1. Estos modelos pertenecen a una clase amplia de *modelos de variables dependientes limitadas*. En éstos, el valor de la variable dependiente está en cierta forma limitado, en vez de ser una variable continua o cuya distribución es observada en su totalidad¹. Los modelos probit y logit también son conocidos como modelos de variables cualitativas o modelos de variables dependientes discretas.

*. Catedrática Auxiliar, Departamento de Economía, Universidad de Puerto Rico, Recinto de Río Piedras.

1. El término “variable continua” se usa aquí en manera laxa. Se refiere a una variable que puede asumir un número grande de valores sin prácticamente ninguna limitación. La variable dependiente “número de empleados en la manufactura” por ejemplo se consideraría continua aunque números decimales de personas no son posible. Tasa de participación toma valores entre 1 y 100 y se consideraría continua.

Existen ciertas variantes del modelo logit tales como el modelo logit multinomial, el modelo logit condicional, el modelo logit ordenado, y el modelo logit anidado (“nested”) multinomial. En estas variantes la variable dependiente asume más de dos valores. En el condicional los atributos de la decisión, en vez de los atributos de la persona que hace la decisión, entran como variables independientes. En ninguno de estos modelos el valor asignado a cada decisión es relevante excepto en el modelo de logit ordenado. En el anidado hay más de una decisión, y estas se toman en secuencia. Por ejemplo, si unirse a una unión laboral, y una vez miembro de esta unión, si acogerse al plan de seguro de salud de la unión.

El modelo probit también tiene sus variantes. De éstas la más usada es el probit bivariable donde dos decisiones se toman en conjunto. Por ejemplo, si participar en la fuerza laboral y si llevar o no los niños a un cuidado; o similar al “nested”, si migrar o no, y para los que migran, si participar en la fuerza laboral o no. Las otras variantes tales como el probit multivariable y el probit ordenado no son usadas frecuentemente y tienden a mostrar problemas de convergencia.

Pocos paquetes estadísticos estiman las formas menos comunes de los modelos probits. Las variantes del modelo logit sin embargo son comúnmente encontradas en paquetes estadísticos. Algunos paquetes estadísticos usan el término “logístico” (logistic) a la par con el modelo logit para distinguir el modelo donde todas las variables independientes son discretas (logístico) de aquél donde las variables independientes incluyen variables discretas y continuas (logit).

III. Problemas con el método de mínimos cuadrados cuando Y es dicótoma:

Cuando la variable dependiente, Y, puede asumir solo dos valores (0,1) el método de mínimos cuadrados presenta ciertos problemas haciéndolo inadecuado para este tipo de estimación:

- No normalidad del término de error: El término de error no es normal, violando uno de los supuestos del modelo clásico de regresión. Si Y es la variable independiente, X son variables explicativas, $X_0 = 1$ representa el término para la constante y μ es el término del error, entonces:

$$\begin{aligned}
 Y &= \sum_{k=0}^{k=K} \beta_k X_k + \mu \\
 \text{Si } Y=1 \quad 1 &= \sum_{k=0}^{k=K} \beta_k X_k + \mu \\
 \mu &= 1 - \sum_{k=0}^{k=K} \beta_k X_k \\
 \text{Si } Y=0 \quad 0 &= \sum_{k=0}^{k=K} \beta_k X_k + \mu \\
 \mu &= - \sum_{k=0}^{k=K} \beta_k X_k
 \end{aligned} \tag{1}$$

Vemos que μ no está distribuido normalmente. De hecho, μ sigue una distribución binomial.² Aún así, se puede invocar el teorema del límite central para usar mínimos cuadrados ordinarios. El teorema del límite central establece que cuando el número de observaciones tiende a infinito el promedio de una variable aleatoria no normal es normal.

- **Heterocedasticidad:** El término de error no tiene varianza constante violando otro de los supuestos del modelo clásico de regresión. Refiriéndonos a la ecuación 1 vemos que el valor esperado de μ y su varianza σ son:

$$\begin{aligned}
 0 &= E(\mu) = P \left(1 - \sum_{k=0}^{k=K} \beta_k X_k \right) + (1-P) \left(- \sum_{k=0}^{k=K} \beta_k X_k \right). \\
 \text{resolviendo para } P \text{ tenemos } P &= \sum_{k=0}^{k=K} \beta_k X_k \text{ y sustituyendo:} \\
 \sigma^2 &= P(1-P)^2 + (1-P)P^2 = P(1-P).
 \end{aligned} \tag{2}$$

De manera que la varianza depende del valor de X y por lo tanto no es constante. Los estimados de los parámetros no son eficientes. El problema de heterocedasticidad en este modelo se puede corregir ajustando las variables Y y X por un factor que incluye la varianza.

2. Una distribución binomial es la distribución de probabilidad de “éxito” (Y=1, o ocurrencia del evento) a partir de n pruebas independientes Bernoulli., donde cada prueba tiene la misma probabilidad de éxito (Wooldridge Jeffrey M. Introductory Econometrics p.772, South Western College Publishing , 2000). Una variable aleatoria que sólo puede tomar valores de 0 y 1 se le llama variable aleatoria Bernoulli. En el caso bajo estudio, Y sería Bernoulli y su distribución de probabilidad implicada por μ , el cual asume dos valores, sería binomial.

- Predicciones fuera de rango: Aún cuando el problema de no normalidad y el de heterocedasticidad pueden ser resueltos, la estimación del modelo de mínimos cuadrados no garantiza que el valor predicho de Y estará entre 0 y 1. Esta es la mayor desventaja del mínimos cuadrados ordinarios cuando la variable dependiente está entre 0 y 1.

IV. Las ecuaciones generadas por los modelos probit y logit

En el modelo probit la variable Y proviene de una variable latente Y*. Digamos que hay una relación

$$Y^* = \beta_0 X_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \mu$$

$$Y^* = \sum_{k=0}^K \beta_k X_k + \mu \quad (3)$$

Donde X son variables explicativas, $X_0=1$ representa la constante, y μ el término de error. Y* no es observada; no hay datos para Y*. Lo que se observa es una variable dicótoma Y que asume un valor de 0 o de 1 dependiendo de si la variable latente Y* cruza o no un límite, por ejemplo, 0. Esto es:

$$Y=1 \quad \text{si} \quad Y^* = \sum_{k=0}^K \beta_k X_k + \mu > 0$$

$$= \mu \geq - \sum_{k=0}^K \beta_k X_k$$

$$Y=0 \quad \text{si} \quad Y^* = \sum_{k=0}^K \beta_k X_k + \mu \leq 0$$

$$= \mu \leq - \sum_{k=0}^K \beta_k X_k \quad (4)$$

Note que μ proviene de una variable Y* que es continua. Entonces se puede presumir μ tiene una distribución normal. El evento $Y=1$ se puede expresar como una probabilidad cuya distribución está dada por la función de distribución acumulativa de la normal estándar como sigue:

$$\begin{aligned}
 Pr(Y=1) &= Pr(\mu > -\sum_{k=0}^K \beta_k X_k) = 1 - F\left(\frac{-\sum_{k=0}^K \beta_k X_k}{\sigma}\right) = \Phi(\mathbf{BX}) \\
 Pr(Y=0) &= Pr(\mu \leq -\sum_{k=0}^K \beta_k X_k) = F\left(\frac{-\sum_{k=0}^K \beta_k X_k}{\sigma}\right) = 1 - \Phi(\mathbf{BX})
 \end{aligned}
 \tag{5}$$

Donde, Φ es la función de distribución acumulativa de la normal estándar. La Y es ahora una probabilidad y los valores de F para cada valor de Y están entre 0 y 1. Recuerde que la función de distribución acumulativa de la normal estándar para $x\beta$ está dada por:

$$P = \Phi(\mathbf{XB}) = \int_{-\infty}^{\mathbf{XB}} \frac{1}{\sqrt{2\pi}} e^{\frac{-\mu^2}{2}} d\mu
 \tag{6}$$

P se puede transformar para expresarse como una función lineal de X. Resolviendo la ecuación 6 para $\beta_k X_k$

$$\Phi^{-1}(P) = \sum_{k=0}^K \beta_k X_k
 \tag{7}$$

Osea que la combinación lineal de las X y los estimados β en un modelo probit produce como variable dependiente la inversa de la función de distribución acumulativa.

El caso de la variable latente y su expresión dicótoma se puede ilustrar con un ejemplo basado en migración. El individuo evalúa los costos y los beneficios de la migración y decide migrar si los beneficios son mayores que los costos. En este caso $Y^* = \text{Beneficios} - \text{costos}$, Y es 1 si migra y 0 si no migra:

$$\begin{aligned}
 Y=1 & \text{ si } Y^* = \text{beneficios} - \text{costos} > 0 \\
 Y=0 & \text{ si } Y^* = \text{beneficios} - \text{costos} \leq 0
 \end{aligned}
 \tag{8}$$

No se observan los costos ni los beneficios. Lo que se observa es una variable dicótoma Y cuyo valor de 0 y de 1 depende de si Y^* es mayor o menor que cero.

En el modelo logit también tenemos una variable dicótoma:

$Y=1$ si una condición se reúne

$Y=0$ si esa condición no se reúne

Este puede verse en la misma forma que el modelo probit con la diferencia de que la función de distribución acumulativa F sería la función logit en vez de la normal. En el modelo logit la variable dicótoma Y no necesariamente esconde una variable continua latente Y^* . La variable Y del modelo logit puede resultar de un proceso de decisión cualitativa. En caso del modelo logit la probabilidad del evento ($Y=1$) sigue una distribución acumulativa logit:

$$P = \frac{e^{\sum_{k=1}^K \beta_k X_k}}{1 + e^{\sum_{k=1}^K \beta_k X_k}} \quad (9)$$

Si $Y=0$ entonces:

$$P = \frac{1}{1 + e^{\sum_{k=0}^K \beta_k X_k}} \quad (10)$$

Si queremos expresar la función anterior como una combinación lineal de β y X entonces tomamos logaritmos a ambos lados. La probabilidad del evento $Y=1$ es:

$$\log\left(\frac{P}{1-P}\right) = \sum_{k=0}^K \beta_k X_k \quad (11)$$

Esta ecuación muestra la razón logarítmica de las probabilidades como una combinación lineal de X y los estimados β . Si sustituimos los valores de β y X la variable dependiente

es la razón logarítmica de las probabilidades. De igual forma tomando el anti-log de la ecuación 11 tenemos:

$$\frac{P}{1-P} = e^{\sum_{k=0}^K \beta_k X_k} \quad (12)$$

Esto es, la combinación lineal exponenciada de los estimados de β y las X da la razón de las probabilidades o la razón de las apuestas (ej: las apuestas son de 3:1 a favor del candidato número 1, o el candidato número 1 es tres veces más probable a ganar que el candidato número 2).

V. ¿Cuál estimar: un probit o un logit?

- Cualitativamente son similares: Los modelos probit y logit producirán resultados cualitativamente similares. De hecho se pueden obtener los coeficientes logit a partir de los coeficientes probit y viceversa. Si el coeficiente probit se multiplica por 1.814 obtenemos un valor aproximado del coeficiente logit. El factor 1.61 también se ha propuesto para obtener el coeficiente logit.
- Preferencia de los economistas: Los economistas tienden a usar más el modelo probit que el logit por dos razones. Primero, el probit implica una variable latente continua en la cual el investigador puede estar interesado. Si la variable latente es de interés entonces el modelo probit puede ser recomendable. En el ejemplo de migración, discutido anteriormente, el investigador puede estar interesado en modelar la probabilidad de migración como función de variables proxies de costos y beneficios. Esta ecuación de costos y beneficios es entonces de interés. La otra razón por la cual el probit es más común que el logit entre economistas es porque el error de la variable dependiente truncada en modelos de auto-selección son generados por una distribución normal y por lo tanto se pueden modelar con un probit. Modelos de auto-selección se han proliferado en la micro-econometría en los últimos años.
- Todas las variables discretas: Si todas las variables independientes son discretas el modelo logit es recomendable debido a que permite una interpretación más clara de los coeficientes.
- Concentración de observaciones en los extremos: En el caso de un número de observaciones grande y una gran concentración de estas observaciones en las colas de la distribución se recomienda el uso del modelo logit. La distribución logit tiene más masa en las colas que la distribución normal.

Observaciones que la distribución normal y por consiguiente el modelo probit asignaría una baja probabilidad o clasificaría como inusuales, serían más comunes bajo el modelo logit. Como muestra la siguiente gráfica, la función de distribución acumulativa del logit y del probit se parecen a la parte izquierda de una campana. La distribución probit toca el eje de X y el tope más rápido que la distribución logit. Esto sugiere que es en las colas donde el logit y el probit difieren.

VI. Estimación del modelo

Cuando la variable dependiente es dicótoma el valor $E(Y/X)$ no es lineal en los parámetros β (ecuaciones 6 y 9). Tenemos que usar métodos no lineales para estimar los parámetros. Usamos el método de máxima verosimilitud (MLE para Maximum Likelihood Estimator) para estimar los parámetros, en vez del método de cuadrados mínimos ordinarios. Recuerde que el método de cuadrados mínimos ordinarios minimiza el error. El método de máxima verosimilitud sin embargo, busca los valores de los estimados de manera que se maximice la probabilidad de que el valor observado de la variable dependiente sea predicha a base de los valores de la variables independientes. Bajo MLE definimos la función logarítmica de probabilidad como:

$$l_i(\beta) = Y_i * \log[G(x_i\beta)] + (1 - Y_i) * \log[1 - G(x_i\beta)]$$

$$L(\beta) = \sum_{i=1}^n l_i(\beta) \quad (13)$$

El estimador de máxima verosimilitud del parámetro β maximiza el valor de la función L. La primera ecuación es el valor para cada observación i ; la segunda ecuación agrega este valor para todas las observaciones. Note que el valor de la primera ecuación depende del valor de Y . Si $Y=1$ el segundo término se elimina; si $Y=0$ el primer término se elimina. Si estimamos un modelo logit, G es una función de distribución acumulativa logit, como la de la ecuación 9 y el estimador MLE estimará un modelo logit. Si estimamos un modelo probit G es una función de distribución acumulativa normal como la de la ecuación 6 y el MLE estimará un modelo probit.

El estimador de máxima verosimilitud tiene propiedades asintóticas deseables. Este estimado es consistente, asintóticamente normal y asintóticamente eficiente. Recuerde que asintótico se refiere a las propiedades de los estimadores en muestras grandes. De manera que las propiedades deseables de los estimados de logit y probit recaen fuertemente en que se tenga una muestra grande. Al menos 100 es aceptable.

VII. Prediciendo el valor de Y :

Para predecir o proyectar el valor de Y sustituimos los valores de los estimados β obtenidos por MLE y el valor promedio de cada variable independiente. Digamos que esto da un valor κ . En el modelo probit buscamos la función de distribución acumulativa de κ . Esto nos da la probabilidad de que $Y=1$. Siempre se predice la probabilidad de $Y=1$ y no de $Y=0$. Paquetes estadísticos como LIMDEP y SAS contienen la función de distribución acumulativa normal en sus comandos.

En el modelo logit sustituimos en la fórmula para P en la ecuación 9. Si se prefiere predecir las razones de probabilidad sustituimos en la ecuación 12. Algunos paquetes estadísticos tienen la función logit y predicen la razón de las probabilidades logarítmicas como en la ecuación 11.

Se puede usar cualquier valor de X que el investigador considere necesario en la predicción. No tiene que usarse el promedio. Si queremos obtener una probabilidad para un caso específico sustituimos esos valores en X. Por ejemplo: si queremos saber la probabilidad de participar en la fuerza laboral para mujeres que tienen 12 años de escolaridad, sustituimos 12 en el valor de X en vez de la escolaridad promedio. Los valores de las otras variables X se pueden tomar en su promedio, o cualquier simulación que se desea. Si la variable independiente es cualitativa, por ejemplo, género, tomando valores de 0 para hombres y de 1 para mujeres, podemos tomar el promedio de la variable discreta aunque ello no tenga mucho significado. Es preferible en este caso sustituir 0 ó 1 y obtener predicciones para hombres o mujeres según sea el caso.

El valor proyectado estará entre 0 y 1 porque es una probabilidad. Si se quiere saber cuantas Y proyectadas son 0 y cuántas son 1 podemos asignar probabilidades mayores de .50 como 1 y aquellas menores o iguales a .50 como 0.

VIII. Contrastes importantes en los modelos probits y logits.

El contraste F usado en cuadrados mínimos ordinarios para contrastar la validez del modelo estimado no puede ser utilizado con probit ni logit pues este supone un modelo lineal. En vez del contraste F usamos el cociente de verosimilitudes logarítmicas (LR para log-likelihood ratio) el cual es:

$$LR = -2(L_0 - L) \approx \chi_r^2 \quad (14)$$

L_0 es el valor de la función de máxima verosimilitud cuando el modelo incluye solamente la constante. L es el valor de la función de máxima verosimilitud con las variables independientes y la constante en el modelo. LR sigue una distribución Chi-cuadrado con r grados de libertad donde r es el número de variables excluidas cuando el modelo se estima sólo con la constante.

L_0 y L son rutinariamente reportados en estimaciones de modelos logit y probit.

Este contraste también se puede usar en general para comparar modelos restringidos y no restringidos tales como aquellos en que se excluye un subconjunto de variables y su ejecutoria se compara con el modelo que incluye todas las variables. En esta situación se usa el valor de la función logarítmica de máxima verosimilitud del modelo restringido en

vez de L_0 , con la del modelo no restringido. La distribución sigue siendo chi-cuadrado con r (número de restricciones) grados de libertad.

El contraste t se usa para hacer pruebas de hipótesis de los coeficientes.

Para obtener una medida de bondad de ajuste se han propuesto varios contrastes. Sin embargo tome en cuenta que bondad de ajuste tiende a ser secundario en este tipo de modelo y que el enfoque tiende a ser significación estadística y económica. El índice del cociente de verosimilitudes: $ICV=1-(L/L_0)$ donde L_0 y L se definen como en la ecuación 14. También se le conoce como R^2 McFadden. Otra alternativa de bondad de ajuste se basa en el porcentaje de observaciones proyectadas correctamente. Se le asignan a aquellas probabilidades proyectadas mayores de .50 un valor de 1 y a las menores o iguales a .5 un valor de cero. Entonces se obtiene la proporción de las observaciones que son proyectadas correctamente.

IX. Interpretación de los coeficientes:

- Signo: Si uno de los coeficientes, digamos β_1 , es mayor que cero decimos que cuando X_1 aumenta, la probabilidad de que Y sea 1 aumenta. Esto es cierto en modelos probits y logits. O sea un coeficiente >0 siempre se asocia con un aumento en la probabilidad de $Y=1$ mientras que un coeficiente negativo siempre se asocia con una baja en la probabilidad de $Y=1$. El signo no nos dice el tamaño del efecto solo su dirección. Si la variable independiente cuyo signo es mayor que cero es dicótoma entonces decimos que la presencia de esa característica se asocia con un aumento en la probabilidad de que $Y=1$. Por ejemplo, 0=hombre; 1=mujer, cuando el coeficiente de esta variable es positivo decimos que la probabilidad de que $Y=1$ es más alta entre las mujeres que los hombres, o que las mujeres tienen una probabilidad más alta de que $Y=1$.
- El efecto marginal: Si sustituimos los valores promedios de X y los estimados de β en la ecuación 7 obtenemos como variable dependiente la inversa de la función de distribución acumulativa en el caso probit. Si sustituimos en la ecuación 11 obtenemos el logaritmo de la razón de las probabilidades en el modelo logit. Aunque esto es correcto, este procedimiento para interpretar el efecto de los estimados no es recomendable pues estas variables dependientes no son inmediatamente intuitivas. No pensamos en términos de la inversa de la función de distribución ni en términos de la razón logarítmica de las probabilidades.
- El efecto marginal en la probabilidad del evento: Este procedimiento usa los estimados de β en forma modificada para obtener algo similar a la pendiente como se entiende en cuadrados mínimos ordinarios. Note que la relación entre X y Y no es lineal y que el valor de la pendiente depende del valor de

X. Se puede obtener un efecto promedio del efecto marginal en la probabilidad de Y cuando X cambia por una unidad. En el modelo probit la “pendiente” de X_k es:

■

$$\Delta P / \Delta X_k = \phi(BX) \beta_k \quad (15)$$

donde ϕ es la función de densidad normal evaluada en los estimados y los valores promedios de X. La función de densidad está disponible en varios paquetes estadísticos como SAS y LIMDEP. El coeficiente para la variable k es multiplicado por la densidad evaluada para obtener el efecto de X en Y. Para estimar el efecto marginal de un coeficiente en la probabilidad de Y en un modelo logit obtenemos:

$$\Delta P / \Delta X_k = P(1-P) \beta_k \quad (16)$$

Donde P es el por ciento de observaciones con Y=1. Este es el método más usado por economistas. Cuando la variable explicativa es dicótoma este procedimiento produce un pequeño error al predecir la probabilidad. En ambos caso decimos que cuando X aumenta por una unidad, la probabilidad de que Y=1 aumenta por el término obtenido en la ecuación de la pendiente. Por ejemplo si el coeficiente ajustado del modelo probit para la variable género (1=mujer) es -.31 decimos que la probabilidad del evento Y=1 es .31 menor para las mujeres que para los hombres.

■

Probabilidad proyectada dado unos valores para X: El problema presentado por las variables dicótomas independientes en el método anterior motiva a que se prefiera el uso de probabilidades proyectadas. En este se sustituye exactamente el valor de X que se quiere investigar y se computa la probabilidad. Si se desea obtener el efecto de un aumento de w unidades en X se puede sustituir el valor base de X y obtener la probabilidad logit o probit y comparar estas probabilidades con el valor nuevo de X. Por ejemplo si se quiere saber el efecto de género en la probabilidad de continuar estudios graduados, estimamos el modelo y usamos el coeficiente de edad para predecir probabilidad. Se computa el valor de la función de distribución acumulativa normal en el caso probit para edad=w1 y para edad=w2. Estas dos probabilidades se comparan y la diferencia es el efecto de edad en la probabilidad de continuar estudios graduados. Digamos que este término es -.33, querría decir que cuando la edad aumenta por un año la probabilidad de participación se reduce por .33.

- El efecto en el cociente de las probabilidades (coeficiente exponencial): En un modelo logit en el cual todas las variables son cualitativas puede ser atractivo usar el coeficiente logit exponencial para obtener el cociente de probabilidades de un evento como se ilustra en la ecuación 12. Por ejemplo si en un modelo estimado la probabilidad de obtener A en el curso de econometría depende de si usted faltó menos de 4 veces a la clase o no, y el coeficiente β de esta variable es .33. Entonces el coeficiente exponencial es $e^{.33}=1.39$ lo cual quiere decir que los estudiantes que faltaron menos de 4 veces a la clase son 39 por ciento más probables a obtener A que los estudiantes que faltaron 5 veces o más. Igualmente se puede decir que los que faltan 4 veces o menos a la clase son 1.39 tan probables a obtener A como los que faltan 5 veces o más. Note que si los que faltan más de 4 veces fueran igualmente probables a obtener A que los que faltan 5 o más veces el coeficiente exponencial sería 1. Si $\beta=-.15$ entonces $e^{-.15}=.86$ diríamos que los estudiantes que faltan 4 veces o menos son .86 tan probables a obtener A como los que faltan 5 veces o más. Igualmente se puede decir que los que faltan 4 veces o menos son 14 por ciento (1-.86) menos probables que los que faltan 5 veces o más a obtener A.

X. Ejemplos de interpretación de coeficientes:

A continuación se presentan varios ejemplos de tomados de la literatura económica para ilustrar el uso y la interpretación de modelos logit y probit.

(1) Ejemplo probit usando pendiente y probabilidades proyectadas: Tomado del artículo "Welfare Payments and Other Economic Determinants of Female Migration," por María E. Enchautegui, *Journal of Labor Economics* 1997 15(3) pt. 1. Variables seleccionadas de la Tabla 3. El modelo predice la probabilidad de cambiar lugar de residencia para madres solteras, edades 25-54. Las cifras en paréntesis son errores estándares.

Variable	Coeficiente	Pendiente
EDAD	-.011(.005)***	-.011*.12=.001
RAZA NEGRA	-.465 (.074)***	-.465*.12=-.05
EDUCACION	.090(.013)***	.090*.12=-.01
ASISTENCIA PÚBLICA (destino-origen)	.027(.002)***	.027*.12=.003

Usando la ecuación 15 se calculó el valor de la función de densidad en los promedios de todas las variables independientes obteniendo un valor de .12. Los coeficientes se multiplican por .12 para obtener la pendiente. Este es el factor de ajuste. En el caso del efecto de edad: un aumento de 1 año en edad está asociado con una baja en la probabilidad de migrar de .001 ($.12 * -.011$). Un aumento de 5 años se asocia con una baja en la probabilidad de migrar de .005 ($.12 * -.011 * 5$). Ser de raza negra está asociado con una baja en la probabilidad de migrar de .05 ($.12 * -.465$).

También se proyectaron las probabilidades. Se calculó el valor de la función de distribución acumulativa (la probabilidad en un modelo probit) en el promedio de todas las variables. Se aumentó el valor de la variable continua de asistencia pública en 10 por ciento partiendo de su promedio. Se re-calculó la probabilidad usando el nuevo valor de asistencia pública re-calculando el valor de la función de distribución acumulativa. Se obtuvo que un aumento de 10% en los beneficios de asistencia pública aumenta la probabilidad de migración de 9.1 a 9.3 por ciento, o por 2 por ciento.

(2) Ejemplo logit usando la pendiente: En el artículo “The absence of African-American Owned Business: An analysis of the Dynamics of Self-employment” de Robert W. Fairlie, *Journal of Labor Economics*, January 1999 Vol 17 (1):80-108, se presentan los siguientes resultados de un modelo logit estimando la probabilidad de estar auto-empleado. Los siguientes resultados son de variables seleccionadas, de la tabla 2: Ecuación logit prediciendo probabilidad de auto-empleo: Hombres negros

Variable	Coeficiente	Pendiente
Casado	-.776 (.307)	$-.776 * .021 = .016$
Divorciado	-.696 (.383)	$-.696 * .021 = .014$
Ingreso concepto interés	.604 (.174)	$-.604 * .021 = .012$
Padre auto-empleado	1.30 (.49)	$1.30 * .021 = .027$

En este ejemplo se usa el efecto marginal transformado o la pendiente para obtener los efectos de las variables independientes en la probabilidad de auto-empleo. El factor de ajuste dado por la ecuación 16 es .021. Tener un padre auto-empleado aumenta la probabilidad de convertirse en auto-empleado por .027. Un aumento de \$1000 en ingreso por concepto de interés aumenta la probabilidad de auto-empleo por .012. Estar casado reduce la probabilidad de convertirse en auto-empleado por .016.

(3) Ejemplo de logit interpretando coeficientes usando la forma exponencial.: “Technological change and the skill acquisition of young workers.” de Ann Bartel y Nachum Sicherman. *Journal of Labor Economics* Octubre 1998 16(4) 718-755. Apéndice B: Modelo logit proyectando la probabilidad de adiestramiento en el empleo para hombres en la manufactura: variables seleccionadas.

Variable	Coficiente	exp(b)
Maestría o más	1.21 (.209)	exp(1.21)=3.35
Compañía grande	.842 (.119)	exp(.842)= 2.32
Casado	.230 (.121)	exp(.230)=1.25
Raza: No-blanco	-.244 (-.145)	exp(-.244)=.783

Personas que tienen un grado de maestría o más son más de tres veces tan probables como aquellos con sólo diploma de escuela superior a ser adiestrados en el empleo. Los trabajadores negros, Latinos o asiáticos son 32 por ciento menos probables que los trabajadores blancos a ser adiestrados en el empleo. Igualmente se podría decir que los trabajadores negros, Latinos o asiáticos son .78 tan probables como los blancos a ser adiestrados en el empleo. Trabajar para una compañía grande aumenta la probabilidad de adiestramiento en el empleo por más del doble. O en otras palabras, trabajadores en compañías grandes son 2.3 veces tan probables como aquellos en compañías pequeñas a ser adiestrados en el empleo.

(4) Ejemplo probit usando la pendiente probabilidad predecida. Este ejemplo se toma de Jeffrey M. Wooldridge, *Introductory econometrics: A modern approach*. South-western Cap 17 p. 538. Se estima la probabilidad de participación laboral de las mujeres usando un modelo probit. Las siguientes son variables seleccionadas de la tabla 17.1

Variable	Coficiente	Pendiente
Educación	.131 (.025)	.131*.391=.05
años de experiencia	.123 (.019)	.123*.391=.048
No. niños edades 0 to-6	-.868 (.119)	-.868*.391=.339
No. niños edades 7-17	.036 (.043)	.036*.391=.014

Tener un niño entre las edades 0 a 6 reduce la probabilidad de participación laboral por .339 mientras un año adicional de experiencia laboral aumenta la probabilidad de participación por .048. Se pudieron haber calculado la distribución acumulativa para ciertas características y compararla cambiando una característica. Si el ingreso es de 20.13, la educación es 12.3, tiene 10.6 años de experiencia y 42.5 años y un niño 0 a 6 la probabilidad proyectada (la función de distribución acumulativa) es .373. Si la mujer tuviera estas mismas características pero no tuviera niños pequeños la probabilidad proyectada de participación sería .707. Entonces decimos que el efecto de tener un niño es $.373-.707=.334$, casi igual a .339 obtenido con la pendiente en vez de probabilidad esperada.

XI. Referencias

A pesar de la proliferación de modelos de variables limitadas en econometría los libros de texto continúan tratando este tema en los últimos capítulos y sus exposiciones tienen a carecer detalle. Aún así los textos más recientes tienen mejores discusiones sobre el temas que textos más viejos y tradicionales.

Lo siguiente es una lista de referencia que en ninguna forma intenta ser exhaustiva con comentarios sobre su accesibilidad para distintos niveles de estudio.

Amemiya T.1985. Advanced Econometrics. Harvard Press. Discusión altamente teórica recomendado para estudiantes a nivel doctoral.

Greene, William. H. LIMDEP 7.0 Manual. Discusión clara y concisa sobre cada modelo de variables cualitativas y limitadas incluido en el paquete econométrico LIMDEP.

Greene. William H. 1999. Análisis econométrico, 3ra edición. Traducido por Hernández Sanchez Antonio, et al. Prentice Hall. Discusión de logit y probit, y otros modelos de variables limitadas. Accesible a estudiantes de maestría.

Liao, Tim Futing. 1994. Interpreting probability models: Logit, Probit, and Other Generalized Linear Models. Series: Quantitative Applications in the Social Sciences no. 101. Sage Publications. Tratamiento básico de modelos probit, logits, y sus variantes. Accesible a todos los niveles

Long J. Scott. 1997. Regression models for categorical and limited dependent variables. Advanced Quantitative Techniques in the Social Sciences no 7. Sage Publications. Un tratamiento avanzado de modelos de variables cualitativas. Recomendable a nivel doctoral.

Maddala G.S. 1983. Limited Dependent and Qualitative Variables in Econometrics. Cambridge University Press. Discusión detallada de todos los modelos de variables dependientes. especialmente de modelos de distribuciones truncadas y censuradas. Algunos capítulos son accesible para estudiantes de bachillerato. Referencia obligada para interesados en la materia.

Ramanathan Ramu. 1999. Introductory Econometrics with Applications. Drysdren Press. Discusión de logit, probit y tobit. Accesible a estudiantes de bachillerato.

Wooldridge Jeffrey M. 2000. Introductory Econometrics: A modern approach. South Western. Discusión de logit y probit y otros modelos de variables limitadas. Accesible a estudiantes de bachillerato.