Microeconometría I

Maestría en Econometría

Lecture 2

Modelos Logit y Probit

1. Temas Avanzados de Logit y Probit
   - Medidas de Diagnóstico
     - Método "informal" para contrastar por error en la especificación de la forma funcional
     - Método formal para contrastar por error en la especificación del argumento de la forma funcional
   - Heterocedasticidad
   - Endogeneidad
     - Variable endógena continua
     - Variable endógena dicotómica

Agenda

1. Temas Avanzados de Logit y Probit
   - Medidas de Diagnóstico
     - Método “informal” para contrastar por error en la especificación de la forma funcional
     - Método formal para contrastar por error en la especificación del argumento de la forma funcional
   - Heterocedasticidad
   - Endogeneidad
     - Variable endógena continua
     - Variable endógena dicotómica

Forma Funcional

- Recuerde que abandonamos el MPL esencialmente porque la forma funcional de la probabilidad no era correcta.
- En su lugar especificamos la probabilidad de ocurrencia del evento analizado con la siguiente función:

$$Pr(y_i = 1|x) = G(\beta_1 + \beta_2 x_2 + \cdots + \beta_K x_K)$$

  donde $G(\cdot)$ es la función de distribución logística (modelo Logit) o la función de distribución normal estándar (modelo Probit).
- Primero, no hay ninguna garantía de que las funciones de distribución de la normal o la logística sean las formas funcionales adecuadas.
- Segundo, es posible que la función asumida para el modelo $\beta_1 + \beta_2 x_2 + \cdots + \beta_K x_K$ pueda ser incorrecta.

Forma Funcional

- Es decir, hay dos fuentes de error en la especificación del modelo:
  - la función $G(\cdot)$, usualmente normal o logística, puede ser incorrecta.
  - el argumento de la función $G(\cdot)$ puede tener una forma funcional incorrecta.
- Necesitamos algunos métodos que nos ayuden a detectar este tipo de errores en la forma funcional.
- Vamos a desarrollar dos de esos métodos:
  - un método "informal" semi-paramétrico.
  - un método formal y paramétrico.

Forma Funcional

- Supongamos que acabamos de estimar el siguiente modelo Probit:

$$Pr(y_i = 1|x) = \Phi(\beta_1 + \beta_2 x_2 + \cdots + \beta_K x_K)$$

  (si es un modelo Logit, solo reemplace $\Phi(\cdot)$ por $F(\cdot)$).
- La hipótesis nula es que la forma funcional es correcta.
- Esto es, la función $\Phi(\cdot)$ es correcta y su argumento $\beta_1 + \beta_2 x_2 + \cdots + \beta_K x_K$ también.

Forma Funcional

- Dada la estimación de los parámetros $\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_K$ calculemos el argumento de la función:

$$\hat{\beta}_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_K x_K = x\hat{\beta}$$

- Bajo la hipótesis nula tenemos:

$$\begin{aligned} Pr(\widehat{y = 1}|x) = \widehat{E(y|x)} &= \Phi(x\hat{\beta}) \\ y &= \Phi(x\hat{\beta}) + \nu \end{aligned}$$

Forma Funcional

- Se sigue que si regresamos y contra $x\hat{\beta}$ permitiendo una forma funcional $\Lambda(\cdot)$ general (i.e. no necesariamente normal):

$$y = \Lambda(x\hat{\beta}) + \nu$$

  la función estimada $\widehat{\Lambda(x\hat{\beta})}$ debería parecerse a la función de distribución de la normal estándar.
- Para permitir una forma funcional flexible para $\Lambda(\cdot)$ se utiliza un procedimiento semi-paramétrico.
- Un método fácil de implementar en Stata es lowless smoothing.
- Una explicación muy informal del procedimiento es como sigue.

Lowless Smoothing

- Para cada observación en los datos calcular el valor esperado de y_i dado x_i desde una regresión que
  1. use solo observaciones de x que estén "cerca" de x_i; y
  2. use ponderadores determinados por la cercanía de cada x_j con x_i (más cerca de x_i, mayor ponderador).
- Una vez que este procedimiento ha sido implementado para todos los datos de la muestra, podemos graficar las n estimaciones de la esperanza matemática de y dado x en el eje vertical y x en el eje horizontal.
- Este gráfico es nuestra estimación de $\Lambda(\cdot)$
- Comparando con el gráfico de $\widehat{\Phi(\cdot)}$ podemos decidir informalmente si la distribución normal estándar es una buena especificación.

Forma Funcional

- Obviamente, existen procedimientos formales para contrastar la forma funcional.
- Estos procedimientos son muy difíciles de implementar empíricamente.
- La referencia clásica es:

  Horowitz, J. L. 1993. "Semiparametric Estimation of a Work-trop Mode Choice Model," *Journal of Econometrics* 58, pp. 49-70.

Forma Funcional

- El punto de partida es el mismo que en el contraste anterior: $x\hat{\beta}$.
- Bajo la hipótesis nula:

$$Pr(y_i = 1|x) = \Phi(x\beta)$$

- Considere la especificación alternativa

$$Pr(y_i = 1|x) = \Phi\left[(x\beta) + \gamma_1(x\beta)^2 + \gamma_2(x\beta)^3\right]$$

- Bajo la hipótesis nula de que el modelo Probit está bien especificado: $\gamma_1 = \gamma_2 = 0$.

Forma Funcional

- Si esta hipótesis nula se rechaza, entonces hay evidencia estadística para creer que el modelo Probit inicial está mal especificado.
- La hipótesis nula se puede contrastar estimando un modelo Probit con y como variable dependiente y con $(x\beta)$, $(x\beta)^2$ y $(x\beta)^3$ como variables explicativas.
- En esta estimación hay que imponer un coeficiente unitario sobre la variable $(x\beta)$ y asegurarse de que el modelo no tenga constante.

Forma Funcional

- Si se rechaza el modelo Probit (Logit), qué se puede hacer?
- Dos opciones:
  - Cambiar la forma funcional de $\Lambda(\cdot)$.
  - Cambiar la forma funcional del argumento de la función. Es muy probable que agregando términos de ordenes superiores en las variables explicativas resuelva el problema.
- Ir a ejemplo con el Stata.

Agenda

1. Temas Avanzados de Logit y Probit
   - Medidas de Diagnóstico
     - Método “informal” para contrastar por error en la especificación de la forma funcional
     - Método formal para contrastar por error en la especificación del argumento de la forma funcional
   - Heterocedasticidad
   - Endogeneidad
     - Variable endógena continua
     - Variable endógena dicotómica

Heterocedasticidad

- Los estimadores de los modelos Logit y Probit no son consistentes bajo heterocedasticidad.
- En presencia de heterocedasticidad la matriz de varianzas y covarianzas de los coeficientes estimados no es apropiada.
- Este es un problema serio ya que la mayoría de las estimaciones de estos modelos se hace con datos de corte transversal donde el problema de la heterocedasticidad es más frecuente.
- Para ilustrar el problema considere el siguiente modelo de variable latente con una sola variable explicativa:

$$y_i^* = \psi_0 + \psi_1 x_{i1} + u_i$$

- Supongamos que el error u_i es heterocedástico.

Heterocedasticidad

- Consideremos una heterocedasticidad multiplicativa:

$$u_i \sim \text{Normal}(0,\, x_{i1}^2)$$

- Recordemos que no se observa la variable latente y_i^*, lo que observamos es,

$$\begin{aligned} y_i &= 1 \quad \text{si} \quad y_i^* > 0 \\ y_i &= 0 \quad \text{si} \quad y_i^* \leq 0 \end{aligned}$$

- Entonces,

$$y_i = 1 \text{ si } \psi_0 + \psi_1 x_{i1} + u_i > 0.$$

Heterocedasticidad

- Siguiendo con nuestro análisis,

$$\begin{aligned} Pr(y_i = 1|x_i) &= Pr(y_i^* > 0|x_i) \\ &= Pr(\psi_0 + \psi_1 x_{i1} + u_i > 0) \\ &= Pr(\psi_0 + \psi_1 x_{i1} + \sqrt{x_{i1}^2} e_i > 0) \end{aligned}$$

  donde $e_i \sim$ Normal(0, 1).
- Por lo tanto,

$$\begin{aligned} Pr(y_i = 1|x_i) &= Pr\left(e_i < -\frac{1}{x_{i1}}(\psi_0 + \psi_1 x_{i1})\right) \\ &= 1 - \Phi\left(-\frac{1}{x_{i1}}(\psi_0 + \psi_1 x_{i1})\right) \\ &= \Phi\left(\frac{1}{x_{i1}}(\psi_0 + \psi_1 x_{i1})\right) \end{aligned}$$

Heterocedasticidad

- Es decir,

$$Pr(y_i = 1|x_i) = \Phi\left(\psi_0 \frac{1}{x_{i1}} + \psi_1\right)$$

- Podemos ver que la presencia de heterocedasticidad ha alterado radicalmente la forma funcional del modelo.
- Dado el modelo subyacente para la variable latente

$$y_i^* = \psi_0 + \psi_1 x_{i1} + u_i$$

- Uno estaría tentado a especificar el modelo Probit como,

$$Pr(y_i = 1|x_i) = \Phi\left(\psi_0 + \psi_1 x_{i1}\right),$$

pero esta no es la especificación correcta en presencia de heterocedasticidad.

Heterocedasticidad

- Pensemos en el efecto marginal de x_{i1}, la especificación correcta es

$$Pr(y_i = 1|x_i) = \Phi\left(\psi_0 \frac{1}{x_{i1}} + \psi_1\right)$$

- El efecto marginal correcto es,

$$\frac{\partial\, Pr(y_i = 1|x_i)}{\partial\, x_{i1}} = \phi\left(\psi_0 \frac{1}{x_{i1}} + \psi_1\right) \times \left(-\psi_0 \left(\frac{1}{x_{i1}}\right)^2\right).$$

- El signo del efecto marginal es el opuesto al signo de ψ_0 (i.e. la constante el el modelo de variable latente) y no depende del signo de ψ_1 (el coeficiente de pendiente en el modelo de variable latente).

Heterocedasticidad

- Del desarrollo anterior se sigue que si ψ_0 y ψ_1 son positivos, el efecto marginal de x_{i1} sobre la probabilidad de ocurrencia del evento analizado tiene el signo opuesto al efecto marginal de x_{i1} en el modelo de variable latente.
- Por supuesto que este último resultado depende crucialmente de la existencia de heterocedasticidad multiplicativa de la forma planteada y por lo tanto el punto anterior no puede verse como un resultado general.
- El punto principal es que si el error del modelo de variable latente es heterocedástico, entonces se altera la forma funcional del Probit.
- Exactamente cómo se altera depende de la forma de la heterocedasticidad.

Heterocedasticidad

- Supongamos que especificamos el Probit incorrectamente como

$$Pr(y_i = 1|x_i) = \Phi\left(\eta_0 + \eta_1 x_{i1}\right)$$

- Será la estimación del coeficiente η_1 una buena estimación de ψ_1 en el modelo de variable latente

$$y_i^* = \psi_0 + \psi_1 x_{i1} + u_i?$$

- La respuesta es no. Y este es un ejemplo de como la presencia de heterocedasticidad lleva a estimadores no consistentes de los parámetros del modelo de variable latente.
- Cómo habría que proceder si creemos que la heterocedasticidad es un problema en nuestro modelo?

Heterocedasticidad

- Una posibilidad es utilizar el comando hetprob del Stata que estima un modelo Probit generalizado:

$$
\begin{aligned}
y^* &= \beta_1 + \beta_2 x_2 + \cdots + \beta_K x_K + e \\
y^* &= x\beta + e
\end{aligned}
$$

donde,

$$\sigma_e^2 = \left[e^{z\gamma}\right]^2,$$

con z un vector de variables (sin constante) que se piense afectan la varianza de e y γ sus correspondientes coeficientes.

Heterocedasticidad

- Entonces,

$$\begin{aligned} Pr(y=1|x,\, z) &= Pr(y^* > 0|x,\, z) \\ &= Pr(x\,\beta + e > 0|x,\, z) \\ &= Pr(x\,\beta + e^{z\,\gamma}u > 0|x,\, z) \end{aligned}$$

donde u sigue una normal estándar (una, valga la redundancia, normalización).

- Por lo tanto,

$$\begin{aligned} Pr(y=1|x,\, z) &= Pr\left(u > \frac{-x\,\beta}{e^{z\,\gamma}}\right) \\ &= 1 - \Phi\left(\frac{-x\,\beta}{e^{z\,\gamma}}\right) \\ &= \Phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right) \end{aligned}$$

Heterocedasticidad

- Por supuesto que si un regresor, x_k está incluido en x y en z el efecto marginal es más complejo,

$$\frac{\partial \, Pr(y=1|x,\, z)}{\partial \, x_k} = \phi \left(\frac{x\,\beta}{e^{z\,\gamma}} \right) \times \left(\frac{\beta_k - (x\beta)\gamma_k}{e^{z\,\gamma}} \right).$$

- Esto muestra que el efecto marginal no tiene necesariamente el signo de β_k.

Heterocedasticidad

- La función de verosimilitud es,

$$L(\hat{\beta}; x,\ z) = \prod_{i=1}^{n} \left[\Phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right)\right]^{y_i} \times \left[1 - \Phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right)\right]^{1-y_i}$$

- Y su logaritmo natural es,

$$l(\hat{\beta}; x,\ z) = \sum_{i=1}^{n} \left[y_i ln\{\Phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right)\} + (1 - y_i) ln\{1 - \Phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right)\}\right]$$

Heterocedasticidad

- Las condiciones de primer orden para la maximización son,

$$\frac{\partial l(\cdot)}{\partial \hat{\beta}} = \sum_{i=1}^{n} \left[\frac{y_i - \Phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right)}{\Phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right)\left[1 - \Phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right)\right]} \phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right) \right] e^{-z\,\gamma} x = 0$$

$$\frac{\partial l(\cdot)}{\partial \hat{\gamma}} = \sum_{i=1}^{n} \left[\frac{y_i - \Phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right)}{\Phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right)\left[1 - \Phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right)\right]} \phi\left(\frac{x\,\beta}{e^{z\,\gamma}}\right) \right] e^{-z\,\gamma} z(-x\beta) = 0$$

- Estas son ecuaciones no lineales en las incógnitas.
- Se puede contrastar por heterocedasticidad con un test LR.

Heterocedasticidad

- H_0 : Homocedasticidad vs. H_1 : Heterocedasticidad

- Estadístico de contraste:

  $$LR = -2 \times [\log L(\text{Probit}) - \log L(\text{HetProb})] \sim \chi^2_q.$$

  Donde q es la dimensión de γ.

- Este contraste aparece es la salida de la estimación del Stata utilizando hetprob.
- Una interpretación alternativa de este contraste sugeriría que la forma funcional del modelo Probit es incorrecta.
- Una posibilidad aquí es incorporar como variables explicativas en el modelo Probit inicial las variables z elevadas al cuadrado.

Agenda

1. Temas Avanzados de Logit y Probit
   - Medidas de Diagnóstico
     - Método "informal" para contrastar por error en la especificación de la forma funcional
     - Método formal para contrastar por error en la especificación del argumento de la forma funcional
   - Heterocedasticidad
   - Endogeneidad
     - Variable endógena continua
     - Variable endógena dicotómica

Endogeneidad

- Todos los problemas provocados por la correlación entre alguna de las variables explicativas y el error de la ecuación que se estudian en los modelos lineales se trasladan a los modelos Probit y Logit.
- Sin embargo, la estimación y la interpretación de los modelos Probit y Logit con variables instrumentales no es completamente directa.
- En principio si se quiere estimar un modelo Probit o Logit con variables endógenas hay que imponer algunos supuestos bastante fuertes.
- Estos supuestos hacen que el único modelo que se puede estimar es el modelo Probit.
- Además, la estimación depende de si la variable endógena en el modelo es continua o dicotómica.

Endogeneidad

- Comencemos ilustrando el caso de una variable potencialmente endógena continua.
- Escribamos el modelo en forma de variable latente:

$$\begin{aligned} y_1^* &= z_1\delta_1 + \alpha_1 y_2 + u_i && (1)\\ y_2 &= z_1\delta_{21} + z_2\delta_{22} + v_2 = z\delta_2 + v_2 && (2)\\ y_1 &= 1[y_1^* > 0] && (3) \end{aligned}$$

  donde (u_1, v_2) tienen distribución bivariada normal con media cero y son independientes de z.
- Las ecuaciones (1) y (3) constituyen el modelo estructural.
- La ecuación (2) es la forma reducida de y_2, que es endógena si u_1 y v_2 están correlacionados.

Endogeneidad

- Note que si la ecuación estructural se especifica como un Logit habria que pensar que tipo de distribución bivariada podrian tener (u_1, v_2).
- Esta distribución al no ser normal bivariada seguramente sea mucho más compleja de analizar y por eso en la literatura se analiza solo el modelo Probit.
- Un segundo punto a tener en cuenta es que si v_2 tiene distribución normal, entonces y_2 también la tiene y no puede comportarse como una variable discreta (i.e. la variable potencialmente endógena es continua).
- Finalmente, necesitamos asumir que la varianza de u_1 es igual a uno para poder identificar δ_1 y α_1 en (1).

Endogeneidad

- Para ver la necesidad del último supuesto, supongamos que $u_1 \sim N(0, \sigma_1^2)$.
- Usando las ecuaciones (1) y (3) tenemos,

$$
\begin{aligned}
Pr(y_1 = 1 | y_2, z_1) &= Pr(y_1^* > 0 | y_2, z_1) \\
&= Pr(u_1 > -z_1\delta_1 - \alpha_1 y_2 | y_2, z_1) \\
&= Pr(u_1 \leq z_1\delta_1 + \alpha_1 y_2 | y_2, z_1) \\
&= \Phi(z_1 \frac{\delta_1}{\sigma_1} + \frac{\alpha_1}{\sigma_1} y_2)
\end{aligned}
$$

- Esto muestra que la varianza de u_1 y los parámetros de pendiente del modelo no pueden identificarse por separado.

Endogeneidad

- Las variables en z_1 y z_2 se asumen exógenas.
- Note que las variables en z_2 son los instrumentos (restricciones de exclusión).
- Esto implica que y_2 es endógena en (1) si y solo si la correlación entre u_1 y v_2 es diferente de cero.
- En este contexto, la estimación del modelo puede hacerce con el procedimiento de dos etapas de Rivers y Vuong (1988).
- Normalidad bivariante de (u_1, v_2) con $Var(u_1) = 1$ implica que,

$$u_1 = \frac{Cov(u_1,\ v_2)}{Var(v_2)} v_2 + e_1 = \theta_1 v_2 + e_1$$

donde e_1 es independiente de z y v_2 (y por lo tanto de y_2).

Endogeneidad

- Note que $var(u_1) = 1$ implica

$$\begin{aligned} Var(e_1) &= 1 - \frac{[Cov(u_1,\, v_2)]^2}{Var(v_2)} \\ &= 1 - \frac{[Cov(u_1,\, v_2)]}{\sqrt{Var(v_2)}\sqrt{1}} \frac{[Cov(u_1,\, v_2)]}{\sqrt{Var(v_2)}\sqrt{1}} \\ &= 1 - \rho_1^2 \end{aligned}$$

donde $\rho_1 = corr(u_1,\, v_2)$.

- Re-escribiendo el modelo de variable latente tenemos

$$y_1^* = z_1\delta_1 + \alpha_1 y_2 + \theta_1 v_2 + e_1$$

y piense que v_2 es una variable sobre la que podemos condicionar en el modelo Probit.

Endogeneidad

- Entonces
$$Pr(y_1 = 1|z, y_2, v_2) = \Phi\left(\frac{z_1\delta_1 + \alpha_1 y_2 + \theta_1 v_2}{\sqrt{1-\rho_1^2}}\right).$$

- Esto es, un Probit sobre z_1, y_2 y v_2 estima consistentemente $\delta_1/\sqrt{1-\rho_1^2}$, $\alpha_1/\sqrt{1-\rho_1^2}$ y $\theta_1/\sqrt{1-\rho_1^2}$.

- Como $\rho_1^2 < 1$ cada uno de los coeficientes estimados es mayor que el coeficiente que se estimaría si y_2 fuera exógeno.

Endogeneidad

- En la práctica no se observa v_2.
- Rivers y Vuong sugieren el siguiente procedimiento de dos etapas.
  1. Estime por MCC una regresión de y_2 sobre z y obtenga los residuos $\hat{v}_2$.
  2. Estime un Probit de y_1 sobre z_1, y_2 y $\hat{v}_2$.
- Una característica de este procedimiento es que si y_2 es exógeno entonces el coeficiente sobre $\hat{v}_2$ es cero.
- Esto sugiere contrastar por exogeneidad comparando $H_0 : \theta_1 = 0$ con $H_1 : \theta_1 \neq 0$ usando el test t usual en la estimación del Probit en el segundo paso.
- Sin embargo los errores estándar y estadísticos t usuales del Probit no son estrictamente válidos debido a la presencia de $\hat{v}_2$.

Endogeneidad

- Una alternativa a este procedimiento de dos etapas es estimar el sistema (1)-(3) simultaneamente.
- Esto requiere escribir la función de verosimilitud de observar y_1 e y_2 y maximizarla.
- A diferencia del procedimiento de Rivers y Vuong, el método de estimación simultánea provee de estimaciones directas de δ_1 y α_1.
- Una ventaja del enfoque simultáneo es que permite calcular directamente los errores estándar de los coeficientes estimados.
- Este procedimiento de estimación simultánea se denomina Probit de variables instrumentales o IV Provit y es la estimación que se obtiene usando el comando ivprobit en Stata.

Endogeneidad

- Consideremos el caso en el que el Probit contiene una variable explicativa binaria endógena.

$$\begin{aligned} y_1 &= 1[z_1\delta_1 + \alpha_1 y_2 + u_1 > 0] &(4)\\ y_2 &= 1[z\delta_2 + v_2 > 0] &(5)\end{aligned}$$

  donde (u_1, v_2) es independiente de z y se distribuye normal bivariada con media cero. Cada componente tiene varianza unitaria y $\rho_1 = corr(u_1, v_2)$.
- Si $\rho_1 \neq 0$ entonces u_1 e y_2 están correlacionados y una estimación Probit de la ecuación (4) nos dará estimadores inconsistentes de δ_1 y α_1.

Endogeneidad

- Al igual que en el caso de una variable endógena continua, la normalización apropiada para poder identificar a los parámetros de pendiente en la ecuación (4) es que $Var(u_1) = 1$.
- Para obtener la distribuci'on conjunta de (y_1, y_2) condicionada en z recuerde que,

$$f(y_1, y_2|z) = f(y_1|y_2, z)f(y_2|z)$$

- Note que,

$$Pr(y_1 = 1|v_2, z) = \Phi[(z_1\delta_1 + \alpha_1 y_2 + \rho_1 v_2)/(1 - \rho_1^2)^{1/2}].$$

- Como $y_2 = 1$ si y solo si $v_2 > -z\delta_2$ para poder calcular los momentos de la distribución truncada necesitamos asumir que v_2 tiene distribución normal estándar y es independiente de z.

Endogeneidad

- Entonces, la función de densidad de v_2 dado que $v_2 > -z\delta_2$ es

$$\phi(v_2)/Pr(v_2 > -z\delta_2) = \phi(v_2)/\Phi(z\delta_2)$$

- Por lo tanto,

$$\begin{aligned} Pr(y_1 = 1|y_2 = 1, z) &= E[Pr(y_1 = 1|v_2, z)|y_2 = 1, z] \\ &= E\{\Phi[(z_1\delta_1 + \alpha_1 y_2 + \rho_1 v_2)/(1 - \rho_1^2)^{1/2}]|y_2 = 1, z\} \end{aligned}$$

$$= \frac{1}{\Phi(z\delta_2)} \int_{-z\delta_2}^{\infty} \Phi[(z_1\delta_1 + \alpha_1 y_2 + \rho_1 \nu_2)/(1 - \rho_1^2)^{1/2}]\phi(\nu_2)d\nu_2$$

- ν_2 en el integral es el argumento de integración.

Endogeneidad

- Obviamente $Pr(y_1 = 0|y_2 = 1, z) = 1 - Pr(y_1 = 1|y_2 = 1, z)$.
- Similarmente $Pr(y_1 = 1|y_2 = 0, z)$ es,

$$\frac{1}{1 - \Phi(z\delta_2)} \int_{-z\delta_2}^{\infty} \Phi[(z_1\delta_1 + \alpha_1 y_2 + \rho_1 \nu_2)/(1 - \rho_1^2)^{1/2}]\phi(\nu_2)d\nu_2$$

- Combinando los cuatro posibles resultados de (y_1, y_2) junto con el modelo Probit para y_2 y tomando logaritmos se obtiene la función de verosimilitud del modelo.
- En la práctica se puede estimar el modelo de las ecuaciones (4)-(5) utilizando un modelo Probit bivariado.
- En Stata el comando es biprobit.