Revista Ingenierías Universidad de Medellín
ISSN: 1692-3324
revistaingenierias@udem.edu.co
Universidad de Medellín
Colombia

Franco Arbeláez, Luis Ceferino; Zuluaga Díaz, Francisco Iván
El problema de condiciones iniciales en la estimación de procesos estocásticos discretos: algunos elementos teóricos
Revista Ingenierías Universidad de Medellín, vol. 5, núm. 8, enero- junio, 2006, pp. 35-41
Universidad de Medellín
Medellín, Colombia

EL PROBLEMA DE CONDICIONES INICIALES EN LA ESTIMACIÓN DE PROCESOS ESTOCÁSTICOS DISCRETOS: ALGUNOS ELEMENTOS TEÓRICOS

Luis Ceferino Franco Arbeláez* y Francisco Iván Zuluaga Díaz**

RECIBIDO: 28/02/2006
ACEPTADO 28/04/2006

RESUMEN

Este artículo considera el problema de condiciones iniciales que surge en la estimación de procesos estocásticos discretos, y se presentan soluciones potencialmente útiles desde el punto de vista computacional.

Palabras clave

Condiciones iniciales, estimación por simulación, datos de panel, efectos inobservados.

ABSTRACT

This paper considers the problem of initial conditions that arise in estimating discrete stochastic process, and potentially useful solutions appear from a computational viewpoint.

Key words

Initial conditions, Simulation Estimation, Panel Data, unobserved effects

INTRODUCCIÓN

En la modelación de datos de panel no lineales dinámicos, el econometrista se enfrenta a una serie de problemáticas entre las cuales se encuentra el problema de condiciones iniciales. Para hacer frente a este problema, generalmente se hace uso de unos supuestos muy restrictivos entre los cuales se encuentran el suponer que la historia premuestral es exógena o, también, que el proceso se asume en equilibrio.

El primer supuesto sólo es válido si las perturbaciones que generan el proceso son serialmente independientes o si verdaderamente un nuevo proceso es observado al inicio del periodo muestral del cual dispone el econometrista. En caso de no ser así, no se pueden considerar a las condiciones iniciales como una variable exógena.

El segundo supuesto resulta inadecuado en el caso en el cual el fenómeno a estudiar esté gobernado por variables variantes en el tiempo, tal como sucede en la mayor parte de la modelación econométrica.

* Profesor del Programa de Ingeniería Financiera de la Universidad de Medellín
** Profesor del Programa de Ingeniería Financiera de la Universidad de Medellín

Entre las soluciones más destacadas al problema de condiciones iniciales se encuentran las de Heckman (1981), Wooldrige (2003) y Keane y Sauer (2005).

A continuación se expone cada una de estas soluciones.

La solución de Heckman

El modelo considerado por Heckman (1981) es un proceso de Markov de primer orden, el cual está definido por la discretización de la variable latente [illegible]

$$y_{it} = \beta_0 + \gamma d_{i,t-1} + \varepsilon_{i,t} \quad i = 1, \ldots, N, \; t = 1, \ldots T \quad (1.1)$$

Donde:

$$d_{it} = \begin{cases} 1 & si \; y_{it} \geq 0 \\ 0 & si \; y_{it} < 0 \end{cases}$$

$$\varepsilon_{it} = \eta_i + u_{it}$$

Las propiedades que se asumen para el término de error son las siguientes:

$$u_{it} \sim \left(0, \sigma_u^2\right), \; \eta_i \sim \left(0, \sigma_\eta^2\right) \; \text{Y} \; Cov\left(\eta_i, u_{it}\right) = 0.$$

Y por último se asume que [illegible] es no estocástica.

Dado lo anterior, la probabilidad de ocurrencia del suceso para el individuo i en el periodo t se puede representar como

$$\begin{aligned} P\left(d_{it} = 1 \middle| d_{i,t-1}, \eta_i\right) &= P\left(\varepsilon_{it} \geq -\beta_0 - \gamma d_{i,t-1} \middle| d_{i,t-1}, \eta_i\right) \\ &= P\left(\eta_i + u_{it} \geq -\beta_0 - \gamma d_{i,t-1} \middle| d_{i,t-1}, \eta_i\right) \\ &= P\left(u_{it} \geq -\beta_0 - \gamma d_{i,t-1} - \eta_i \middle| d_{i,t-1}, \eta_i\right) \\ &= 1 - P\left(u_{it} \leq -\beta_0 - \gamma d_{i,t-1} - \eta_i \middle| d_{i,t-1}, \eta_i\right) \\ &= 1 - \Phi\left(-\beta_0 - \gamma d_{i,t-1} - \eta_i\right) \\ &= \Phi\left(\beta_0 + \gamma d_{i,t-1} + \eta_i\right) \end{aligned}$$

De esta manera la probabilidad de transición para el individuo i en el periodo t estará dada por

$$P\left(d_{it} \middle| d_{i,t-1}, \eta_i\right) = \Phi\left[\left(\beta_0 + \gamma d_{i,t-1} + \eta_i\right)\left(2d_{i,t} - 1\right)\right] \quad (1.2)$$

La expresión $\left(2d_{it} - 1\right)$ permite tener en cuenta todas las posibles alternativas, donde [illegible] es la distribución normal estándar.

La probabilidad marginal de d_{iT}, dado [illegible] para $T \geq J$, se puede representar como la suma de las probabilidades de todas las posibles sucesiones de los eventos anteriores a J, esto es

$$\begin{aligned} P\left(d_{iJ} \middle| \eta_i\right) = &\left(\sum_{d_{i,J-1}=0}^{1} \Phi\left[\left(\beta_0 + \gamma d_{i,J-1} + \eta_i\right)\left(2d_{ij} - 1\right)\right]\right) \\ &* \left(\sum_{d_{i,J-2}=0}^{1} \Phi\left[\left(\beta_0 + \gamma d_{i,J-2} + \eta_i\right)\left(2d_{i,J-1} - 1\right)\right]\right) \\ &* \cdots * \left(\sum_{d_{i1}=0}^{1} \Phi\left[\left(\beta_0 + \gamma d_{i0} + \eta_i\right)\left(2d_{i1} - 1\right)\right]\right) \end{aligned} \quad (1.3)$$

Asumiendo que el econometrista tiene acceso a toda la historia del proceso, se puede plantear la función de verosimilitud obteniendo primero la función de densidad conjunta para los u_{it}, lo cual puede hacerse partiendo de la densidad conjunta de [illegible] e integrando con respecto a la última variable, dado que $f\left(u_{i1}, \ldots, u_{iT}, \eta_i\right) = f\left(u_{i1}, \ldots, u_{iT} \middle| \eta_i\right) f\left(\eta_i\right)$. Por lo tanto,

$$f\left(u_{i1}, \ldots, u_{iT}\right) = \int_{-\infty}^{\infty} f\left(u_{i1}, \ldots, u_{iT} \middle| \eta_i\right) f\left(\eta_i\right) d\eta_i$$

Ahora dada la independencia de los u_{it} esta expresión se transforma en

$$f\left(u_{i1}, \ldots u_{iT}\right) = \int_{-\infty}^{\infty} \prod_{t=1}^{T} f\left(u_{it} \middle| \eta_i\right) f\left(\eta_i\right) d\eta_i$$

De esta manera la función de verosimilitud para una muestra de T observaciones por individuo dado d_{i0}, es

$$L=\prod_{i=1}^{N}\int_{-\infty}^{\infty}\prod_{t=1}^{T}\Phi\left[\left(\beta_0+\gamma d_{i,t-1}+\eta_i\right)\left(2d_{it}-1\right)\right]f\left(\eta_i\right)d\eta_i \quad (1.4)$$

Como argumenta Heckman (1981), los estimadores de máxima verosimilitud para [illegible] son consistentes cuando $T/N \rightarrow c$ (donde c es una constante positiva).

Cuando el econometrista solamente tenga acceso a las últimas T - J observaciones del proceso, no puede asumir que el estado inicial [illegible] es fijo o exógeno; por tanto, éste resulta ser estocásticamente dependiente de η_i. En este caso la función de verosimilitud para [illegible] dado d_{iT} estará dada por

$$L=\prod_{i=1}^{N}\left\{\int_{-\infty}^{\infty}\prod_{t=J+1}^{T}\Phi\left[\left(\beta_0+\gamma d_{i,t-1}+\eta_i\right)\left(2d_{it}-1\right)\right]*P\left(d_{iJ}|\eta_i\right)f\left(\eta_i\right)d\eta_i \Big/ \int_{-\infty}^{\infty}P\left(d_{iJ}|\eta_i\right)f\left(\eta_i\right)d\eta_i\right\} \quad (1.5)$$

La maximización de (1.5) producirá estimaciones consistentes para los parámetros si $T/N \rightarrow c$; lo que sucede es que la maximización de (1.5) puede ser computacionalmente complicada. Para afrontar esta complicación Heckman (1981) sugiere asumir estacionariedad inicial del proceso. Sin embargo este supuesto es cuestionable cuando variables exógenas sean incluidas en el modelo, lo cual obligaría a suponer que éstas son generadas por un proceso estocástico estacionario, y en caso tal de que no se haga el supuesto de estacionariedad, para poder utilizar (1.4) se hace necesario determinar [illegible], lo cual es complicado; y, a su vez, conocer los valores de las variables exógenas en el período premuestral es un problema difícil.

Finalmente, para dar solución al problema de condiciones iniciales, Heckman (1981) propone utilizar simulación de Monte Carlo. La idea es simplemente aproximar la probabilidad marginal del estado inicial por una función probit que tiene como argumento tanta información premuestral sobre las variables exógenas como sea posible, de la siguiente manera:

1. Aproxímese la probabilidad del estado inicial d_{iT} por una función probit

$$y_{iJ}=\sum_{\ell=0}^{J}D(\ell)x_{it}+\upsilon_0$$

$$d_{iJ}=\begin{cases}1 & si\ y_{iJ}\geq 0\\ 0 & si\ y_{iJ}<0\end{cases}$$

   Donde $D(\ell)x_{it}$ es una función general del vector de variables exógenas [illegible] y $\upsilon_{ij}\sim N(0,\sigma_u^2)$.

2. Permítase que [illegible] este libremente correlacionada con $\varepsilon_{it},\ t=J+1,\ldots,T$.

3. Estímese el modelo por el método de máxima verosimilitud sin imponer ningún tipo de restricción sobre los parámetros

La solución de Wooldridge

Wooldridge (2003) propone como una solución alternativa a la de Heckman (1981) para el problema de condiciones iniciales: modelar la distribución de los efectos individuales condicional al valor inicial del proceso y al conjunto de variables exógenas, argumentando que esta aproximación resulta más flexible y computacionalmente más simple. Siguiendo con la estructura del numeral dos, uno de los modelos presentados por Wooldridge (2003) es el siguiente:

$$P\left(d_{it}=1\left|d_{i,t-1},\ldots,d_{i0},x_i,\eta_i\right.\right)=\Phi\left(x_{it}\beta+\gamma d_{i,t-1}+\eta_i\right) \quad (1.6)$$

$$\eta_i\left|d_{i0},x_i\right.\sim N\left(\alpha_0+\alpha_1 d_{i0}+x_i\alpha_2,\sigma_\eta^2\right) \quad (1.7)$$

[illegible] representa un vector de variables estrictamente exógenas, y a su vez [illegible] es un vector fila de todas las variables explicativas en todos los períodos temporales.

Dados (1.6) y (1.7) la función de densidad condicional se puede representar como

$$f(d_{i1},...d_{iT}|d_{i0},x_{it},\eta_i;\delta)=\prod_{t=1}^{T}\Phi\left[(x_{it}\beta+\gamma d_{i,t-1}+\eta_i)(2d_{it}-1)\right] \quad (1.8)$$

Donde [illegible]

Wooldridge (2003) sugiere que en vez de integrar (1.8) con respecto a [illegible], primero se debe especificar una ecuación concreta para [illegible] así:

[illegible] (1.9)

Donde [illegible].ahora utilizando la variable latente [illegible] el modelo (1.6)-(1.7) se puede replantear como sigue:

$$y_{it}=x_{it}\beta+\gamma d_{i,t-1}+\alpha_0+\alpha_1 d_{i0}+x_i\alpha_2+a_i+u_{it} \quad (1.10)$$

[illegible]

Donde [illegible]. Dado lo anterior, la probabilidad de ocurrencia del suceso para el individuo i en el período t se puede representar como:

$$P(d_{it}=1|d_{i,t-1},...,d_{i0},x_i,a_i)=\Phi(x_{it}\beta+\gamma d_{i,t-1}+\alpha_0+\alpha_1 d_{i0}+x_i\alpha_2+a_i) \quad (1.11)$$

Con lo cual se tiene que:

$$f(d_{i1},...d_{iT}|d_{i0},x_i,a_i;\delta)=\prod_{t=1}^{T}\Phi\left[(x_{it}\beta+\gamma d_{i,t-1}+\alpha_0+\alpha_1 d_{i0}+x_i\alpha_2+a_i)(2d_{it}-1)\right] \quad (1.12)$$

E integrando (1.12) con respecto a [illegible]

$$f(d_{i1},...d_{iT}|d_{i0},x_i;\delta)=\int_{-\infty}^{+\infty}\left\{\prod_{t=1}^{T}\Phi\left[(x_{it}\beta+\gamma d_{i,t-1}+\alpha_0+\alpha_1 d_{i0}+x_i\alpha_2+a_i)(2d_{it}-1)\right]f(a_i)da_i\right\}$$

(1.12)

La diferencia con el modelo estándar es la introducción de [illegible] como variables explicativas adicionales. En este caso Wooldridge (2003) sugiere utilizar (1.13) para estimar [illegible].

La solución de Keane y Sauer

Más recientemente Keane y Sauer (2005) (a quienes denotaremos por KS) proponen una nueva metodología para dar solución al problema de condiciones iniciales, y al problema de observaciones faltantes durante el período muestral, sin la necesidad de utilizar probabilidades condicionales. La idea del algoritmo propuesto por KS, consiste en construir la función de verosimilitud usando solamente simulaciones incondicionales, lo cual se puede hacer asumiendo que las elecciones reportadas están sujetas a errores de medición. Asumiendo tasas de error por clasificación en las elecciones reportadas, se elimina la necesidad de condicionar sobre la historia pasada.

El modelo considerado por KS es el siguiente:

$$y_{it}=\beta_0+\beta_1 x_{it}+\sum_{\tau=0}^{t-1}d_{i\tau}\rho_\tau+\varepsilon_{it} \quad (1.13)$$

donde [illegible] es una variable latente, [illegible] es una variable estrictamente exógena y

[illegible] (1.14)

Obsérvese que la especificación dada por (1.14) es más general que las presentadas en los numerales anteriores ya que incluye la historia total de las pasadas elecciones, el parámetro de ponderación [illegible] permite capturar la importancia de elecciones pasadas. Además se asume que [illegible]. KS asumen que [illegible] está correlacionado serialmente, para lo cual asumen dos fuentes de correlación. La primera de ellas tiene que ver con la existencia de un efecto individual, lo que da origen a un modelo de componentes de error, esto es:

[illegible] (1.15)

Donde [illegible]. La correlación serial puede también derivarse de un proceso AR(1),

[illegible] (1.16)

Denotando a [illegible] como la elección reportada, KS consideran que el modelo general es caracterizado por cuatro tasas de error por clasificación

$$
\begin{aligned}
\pi_{11t} &= P\left(d_{it}^{*}=1 \middle| d_{it}=1\right)\\
\pi_{01t} &= P\left(d_{it}^{*}=1 \middle| d_{it}=0\right)\\
\pi_{00t} &= 1-P\left(d_{it}^{*}=1 \middle| d_{it}=0\right)\\
\pi_{10t} &= 1-P\left(d_{it}^{*}=1 \middle| d_{it}=1\right)
\end{aligned}
\quad (1.17)
$$

Donde [illegible] es la probabilidad de que la primera opción reportada sea elegida [illegible], dado que la primera opción es la verdadera elección [illegible]; es la probabilidad de que la primera opción reportada sea elegida [illegible] dado que la segunda opción es la verdadera elección [illegible] es la probabilidad de que la segunda opción reportada sea la elegida [illegible], dado que la segunda opción es la verdadera elección [illegible] es la probabilidad de que la segunda opción reportada sea la elegida [illegible], dado que la primera opción es la verdadera elección [illegible].

KS asumen dos diferentes tipos de especificaciones para las tasas de error por clasificación. El primer supuesto es que las tasas de clasificación por error son insesgadas, es decir, [illegible], lo que significa que la probabilidad de que un individuo sea observado eligiendo una opción, sea igual a la verdadera probabilidad de que el individuo elija esa opción. Este tipo de clasificación se puede obtener asumiendo que las tasas de error por clasificación, descritas en (1.18) son funciones lineales de la probabilidad de elección verdadera, esto es,

$$
\begin{aligned}
P\left(d_{it}^{*}=1 \middle| d_{it}=1\right) &= E+(1-E)P\left(d_{it}=1\right)\\
P\left(d_{it}^{*}=1 \middle| d_{it}=0\right) &= (1-E)P\left(d_{it}=1\right)
\end{aligned}
\quad (1.18)
$$

Donde E representa la probabilidad de clasificar correctamente los eventos con una baja probabilidad de ocurrencia; de esta manera las tasas de error por clasificación pueden ser escritas como

$$
\begin{aligned}
\pi_{11t} &= E+(1-E)P\left(d_{it}=1\right)\\
\pi_{01t} &= (1-E)P\left(d_{it}=1\right)
\end{aligned}
\quad (1.19)
$$

El segundo tipo de clasificación propuesto por KS no impone la relación lineal dada por (1.19), por lo tanto esto dará origen a un proceso de error por clasificación sesgado en el cual [illegible]. El esquema propuesto se basa en la siguiente función indicadora:

$$
\begin{aligned}
l_{it} &= \gamma_0 + \gamma_1 d_{it} + \gamma_2 d^*_{i,t-1} + \omega_{it} \\
d^*_{it} &= \begin{cases} 1 & si\ l_{it} \geq 0 \\ 0 & si\ l_{it} < 0 \end{cases}
\end{aligned}
\tag{1.20}
$$

Donde [illegible] es una variable latente, [illegible] es la elección reportada y [illegible] es un término estocástico. La característica fundamental de (1.21) es la aparición de [illegible], lo que implica que mientras mayor sea el valor de [illegible] mayor será la persistencia en el tiempo de realizar reportes equivocados. Si se asume que [illegible] se distribuye logísticamente las tasas de error por clasificación se podrán representar como sigue

$$
\begin{aligned}
\pi_{11t} &= \frac{e^{\gamma_0 + \gamma_1 + \gamma_2 d^*_{i,t-1}}}{1 + e^{\gamma_0 + \gamma_1 + \gamma_2 d^*_{i,t-1}}} \\
\pi_{01t} &= \frac{e^{\gamma_0 + \gamma_2 d^*_{i,t-1}}}{1 + e^{\gamma_0 + \gamma_2 d^*_{i,t-1}}}
\end{aligned}
\tag{1.21}
$$

Dados los elementos anteriores, el algoritmo de Máxima Verosimilitud Simulado, propuesto por KS es el siguiente:

1. Para cada individuo *i*, obténgase *M* sucesiones de errores de la distribución conjunta de [illegible] para formar [illegible]
2. Dado [illegible] y la sucesión de errores obtenida en (1), constrúyase *M* historias de elección simuladas [illegible] para cada individuo *i*, de acuerdo a (1.14) y (1.15).
3. Constrúyanse las tasas de error por clasificación [illegible] para cada individuo *i*, donde *j* es la elección simulada y *k* denota la elección reportada.
4. Fórmese un simulador insesgado para la función de verosimilitud para el individuo *i* como sigue:

$$
\hat{P}\left(d^*_{i1}, d^*_{i2}, \ldots d^*_{iT} \middle| \theta, x_i\right) = \frac{1}{M} \sum_{m=1}^{M} \prod_{t=1}^{T} \left(\sum_{j=0}^{1} \sum_{k=0}^{1} \hat{\pi}^m_{jkt} I[d^m_{it} = j, d^*_{it} = k] \right)^{I\left(d^*_{it} \text{ es observada}\right)}
$$

donde [illegible] es el vector de parámetros del modelo y [illegible] es una función indicadora que es igual a uno si [illegible] es observada, y cero en cualquier otro caso.

Conclusión

Como se puede observar, la temática discutida en los renglones anteriores se convierte en un campo de investigación promisorio, que en nuestro medio ha tenido poca difusión tanto en el aspecto teórico, como en la aplicación empírica. Una investigación concreta, por ejemplo en el campo de los modelos de datos de panel no lineales dinámicos, consistirá en los siguientes puntos:

- Mediante un diseño experimental analizar las propiedades en muestras finitas de los diferentes estimadores propuestos.
- Realización de casos de aplicación donde se investiguen entre otras temáticas: la dinámica de las decisiones cualitativas de las firmas y Análisis Intertemporal de la participación de la fuerza laboral.
- Ampliación de las metodologías generalmente utilizadas, dando la posibilidad de incluir variables predeterminadas.

BIBLIOGRAFÍA

ARELLANO, M. 2003. *Discrete choices with Panel Data*. Investigaciones Económicas XXVII (3) 423-458.

ARELLANO, M y B. HONORE. 2001. *Panel Data Models: Some Recent Developments.* In Heckman, J.J. and E. Leamer (eds.), Handbook of Econometrics, vol. 5,chapter 53,North-Holland.

HSIAO, C. 2003. *Analysis of Panel Data*. Second Edition. Cambridge University Press.

HECKMAN, J.J. 1981. *The incidental parameters problem and the problem of initial conditions in estimation a discrete time-discrete data stochastic process*. In: Manski, C.f.

KEANE, MP Y R SAUER. 2005. A *Computationally Practical Simulation Estimation Algorithm for Dynamic Panel Data Models With Unobserved Endogenous State*. Yale University manuscript.

WOOLDRIDGE, J. M. 2003. S*imple Solutions to the Initial Conditions Problem in Dynamic Nonlinear Panel Data Models With Unobserved Heterogeneity.* Journal of Applied Econometrics.

ZULUAGA, F. (2005) "*Econometría de Datos de Panel: Revisión y una aplicación*", Tesis de Maestría, Universidad EAFIT.