# Regresja przełącznikowa

Jerzy Mycielski

Październik 2013

# Regresja przełącznikowa - forma modelu

## Próby przekrojowe

- Równania regresji

$$
\begin{aligned}
y_i &= \mathbf{x}_{1i}\boldsymbol{\beta}_1 + u_{1i} \quad \text{dla} \quad s_i = 1 \qquad \text{(Rezim 1)} \\
y_i &= \mathbf{x}_{2i}\boldsymbol{\beta}_2 + u_{2i} \quad \text{dla} \quad s_i = 0 \qquad \text{(Rezim 2)}
\end{aligned}
$$

- Równanie selekcji - model dla zmiennej binarnej

$$
\begin{aligned}
s_i &= 1 \quad \text{dla} \quad \mathbf{z}_i\gamma \geq u_i \\
s_i &= 0 \quad \text{dla} \quad \mathbf{z}_i\gamma < u_i
\end{aligned}
$$

- Identyfikacja: wśród zmiennych zawartych w równaniu selekcji powinny się znaleźć zmienne, których nie ma w równaniach regresji

# Regresja przełącznikowa - założenia o błądzie losowym

- Zakładamy, że poszczególne błędy losowe są *IID* - nie występuje więc autokorelacja i rozkłady są stałe w czasie
- Macierz wariancji i kowarinacji błędu losowego

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{1u} \\ & \sigma_{22} & \sigma_{2u} \\ & & 1 \end{bmatrix}$$

- Zazwyczaj konieczne jest założenie o normalności łącznego rozkładu błędu losowego

- Zakładamy normalność błędów losowych
- Podstawowy problem

$$
\begin{aligned}
\mathrm{E}\left( u_{i1} \middle| \mathbf{z}_i\gamma \geq u_i \right) &= -\sigma_{1u} \frac{\phi\left(\mathbf{z}_i\gamma\right)}{\Phi\left(\mathbf{z}_i\gamma\right)} \\
\mathrm{E}\left( u_{i2} \middle| \mathbf{z}_i\gamma < u_i \right) &= \sigma_{2u} \frac{\phi\left(\mathbf{z}_i\gamma\right)}{\Phi\left(\mathbf{z}_i\gamma\right)}
\end{aligned}
$$

- Wniosek: jeśli $\mathbf{z}_i$ zawiera część zmiennych zawartych w $\mathbf{x}_{1i}$ i $\mathbf{x}_{2i}$ to nie jest spełniony warunek, że $\mathrm{E}\left( u_{1i} \middle| \mathbf{x}_{1i} \right) = 0$ lub $\mathrm{E}\left( u_{2i} \middle| \mathbf{x}_{2i} \right) = 0$.

- Postać funkcji wiarygodności:

$$
\begin{aligned}
& L\left(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \sigma_{11}, \sigma_{22}, \sigma_{12}, \sigma_{1u}, \sigma_{2u}\right) \\
= & \prod_{i=1}^{N}\left[\int_{-\infty}^{\mathbf{z}_i\boldsymbol{\beta}} f\left(y_i - \mathbf{x}_{1i}\boldsymbol{\beta}_1 \mid u_i\right) du_i\right]^{s_i} \\
& \times \left[\int_{\mathbf{z}_i\boldsymbol{\beta}}^{\infty} f\left(y_i - \mathbf{x}_{2i}\boldsymbol{\beta}_1 \mid u_i\right) du_i\right]^{s_i - 1}
\end{aligned}
$$

# Regresja przełącznikowa - przypadki szczególne

- Tobit
- Model Heckmana
- Estymacja za pomocą *MNW* może się okazać kłopotliwa
- W niektórych szczególnych przypadkach możliwa jest estymacja dwustopniowa

# Definicja łańcucha Markowa

- Wartości zmiennych losowych $X_n$ należą do skończonego zbioru $\{s_1, s_2, \ldots, s_r\}$
- Skończony łańcuch Markowa

$$\Pr\left(X_n = j \middle| X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}\right) = P\left(X_n = j \middle| X_{n-1} = i\right)$$

- Proces ma więc krótką pamięć: prawdopodobieństwo przejścia do innego stanu zależy jedynie od stanu w poprzednim okresie
- Skończony jednorodny łańcuch Markowa jest to skończony łańcuch Markowa o następującej własności:

$$P\left(X_n = j \middle| X_{n-1} = i\right) = P\left(X_{n+s} = j \middle| X_{n+s-1} = i\right) = p_{ij}$$

- Dla jednorodnego łańcucha Markowa prawdopodobieństwa przejścia między stanami nie zmieniają się w czasie

# Macierz prawdopodobieństw przejścia

- Macierzą prawdopodobieństw ma następującą postać:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & \vdots & & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix}$$

- $p_{ij}$ muszą spełniać następujące warunki

$$\sum_{j=1}^{r} p_{ij} = 1$$

$$p_{ij} \geq 0$$

# MSM - forma modelu

- Równanie regresji

$$\begin{aligned} y_t &= \mathbf{x}_t \boldsymbol{\beta}_1 + \sigma_1 \varepsilon_t \quad \text{dla} \quad s_t = 1 & \text{(Rezim 1)} \\ y_t &= \mathbf{x}_t \boldsymbol{\beta}_2 + \sigma_2 \varepsilon_t \quad \text{dla} \quad s_t = 0 & \text{(Rezim 2)} \end{aligned}$$

- Równanie selekcji - łańcuch Markowa ze zmiennymi prawdopodobieństwami przejść (niejednorodny)

$$P\left( s_t = i \right| s_{t-1} = j, \mathbf{z}_t ) = p_{ij} \left( \mathbf{z}_t \right)$$

# MSM - Przełączanie

- Zakładamy, że

$$
\begin{aligned}
s_t &= 1 \quad \text{dla} \quad \mathbf{z}_i \gamma_{s_{t-1}} \geq u_t \\
s_t &= 0 \quad \text{dla} \quad \mathbf{z}_i \gamma_{s_{t-1}} < u_t
\end{aligned}
$$

a więc

$$
\begin{aligned}
p_{1 s_{t-1}}(\mathbf{z}_t) &= 1 - \Phi\left(\mathbf{z}_i \gamma_{s_{t-1}}\right) \\
p_{0 s_{t-1}}(\mathbf{z}_t) &= \Phi\left(\mathbf{z}_i \gamma_{s_{t-1}}\right)
\end{aligned}
$$

# MSM - przełączanie c.d.

- Ponadto zakładamy, że rozkład bezwarunkowy łańcucha Markowa jest stacjonarny:

$$P\left(s_t = i \middle| \bar{\mathbf{z}}\right) = P\left(s = i \middle| \bar{\mathbf{z}}\right) \text{ dla każdego } t$$

$$\begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix} \sim N\left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

- Zakłada się, że stan procesu $i$ jest *nieobserwowalny*

# MSM - estymacja

- Estymacji dokonujemy *MNW*
- W ogólnym przypadku postać funkcji wiarygodności jest dosyć skomplikowana
- Szczegóne przypadki
    - przełączenia egzogeniczne: $\boldsymbol{\rho} = 0$
    - czysty proces Markowa $\gamma_1 = \gamma_2 = \mathbf{0}$

- Łączna funkcja gęstości dla całej próby

$$L(\boldsymbol{\theta}) = \prod_{t=1}^{T} f\left(y_t \middle| \boldsymbol{\Omega}_t, \boldsymbol{\xi}_{t-1}; \boldsymbol{\theta}\right)$$

gdie $\boldsymbol{\xi}_t = (y_t, y_{t-1}, \ldots, y_1)$,
$\boldsymbol{\Omega}_t = (\mathbf{x}_t, \mathbf{x}_{t-1}, \ldots, \mathbf{x}_1, \mathbf{z}_t, \mathbf{z}_{t-1}, \ldots, \mathbf{z}_1)$, $\boldsymbol{\theta} = (\boldsymbol{\beta}_1, \sigma_1, \boldsymbol{\beta}_2, \sigma_2)$.

- Funkcja gęstości dla pojedynczego zdarzenia

$$\begin{aligned} & f\left(y_t \middle| \boldsymbol{\Omega}_t, \boldsymbol{\xi}_{t-1}; \boldsymbol{\theta}\right) \\ = & \sum_i \sum_j f\left(y_t \middle| s_t = i, s_{t-1} = j, \boldsymbol{\Omega}_t, \boldsymbol{\xi}_{t-1}; \boldsymbol{\theta}\right) \times \\ & \Pr\left(s_t = i, s_{t-1} = j \middle| \boldsymbol{\Omega}_t, \boldsymbol{\xi}_{t-1}; \boldsymbol{\theta}\right) \end{aligned}$$

- Prawdopodbieństwo przejścia dla danej obserwacji liczymy rekursywnie przy wykorzystaniu twierdzenia Bayesa oraz tego, że

$$P\left(s_0 = i \middle| \bar{\mathbf{z}}\right) = P\left(s = i \middle| \bar{\mathbf{z}}\right)$$

- Tak więc:

$$\Pr\left(s_t = i, s_{t-1} = j \middle| \boldsymbol{\Omega}_t, \boldsymbol{\xi}_{t-1}; \boldsymbol{\theta}\right) = p_{ij}\left(z_t\right) \Pr\left(s_{t-1} = j \middle| \boldsymbol{\Omega}_t, \boldsymbol{\xi}_{t-1}; \boldsymbol{\theta}\right)$$

- Bezwarunkowe prawdopodobieństwo przejścia liczymy z twierdzenia Bayesa:

$$
\begin{aligned}
& \Pr\left(s_t = i \middle| \boldsymbol{\Omega}_{t+1}, \boldsymbol{\xi}_t; \boldsymbol{\theta}\right) \\
& \Pr\left(s_t = i \middle| \boldsymbol{\Omega}_{t}, \boldsymbol{\xi}_t; \boldsymbol{\theta}\right) \\
= \quad & \frac{f\left(y_t \middle| s_t = i, \boldsymbol{\Omega}_t, \boldsymbol{\xi}_{t-1}; \boldsymbol{\theta}\right)}{f\left(y_t \middle| \boldsymbol{\Omega}_t, \boldsymbol{\xi}_{t-1}; \boldsymbol{\theta}\right)} \\
= \quad & \frac{1}{f\left(y_t \middle| \boldsymbol{\Omega}_t, \boldsymbol{\xi}_{t-1}; \boldsymbol{\theta}\right)} \times \\
& \sum_j f\left(y_t \middle| s_t = i, s_{t-1} = j, \boldsymbol{\Omega}_t, \boldsymbol{\xi}_{t-1}; \boldsymbol{\theta}\right) \times \\
& \Pr\left(s_t = i, s_{t-1} = j \middle| \boldsymbol{\Omega}_t, \boldsymbol{\xi}_{t-1}; \boldsymbol{\theta}\right)
\end{aligned}
$$

- w przypadku, gdy $\boldsymbol{\rho} = 0$

$$
f\left(y_t \middle| s_t = i, s_{t-1} = j, \boldsymbol{\Omega}_t, \boldsymbol{\xi}_{t-1}; \boldsymbol{\theta}\right) = \frac{1}{\sigma_i} \boldsymbol{\phi}\left(\frac{y_t - \mathbf{x}_t \boldsymbol{\beta}}{\sigma_i}\right)
$$