

1 Heckman modelis

Šiai užduočiai bus naudojami duomenys apie ištekėjusių moterų dalyvavimą darbo rinkoje. Duomenys yra R pakete „micEcon“, juos gausite suinstaliavę tą paketą bei parašę `library(micEcon)` ir `data(Mroz87)` R komandinėje eilutėje. Duomenis taip pat bus galima atsisiųsti iš puslapio, kuriame patalpinta šita užduotis.

1. Atlikite pradinę duomenų analizę.
2. Raskite geriausią probit modelį nusakantį moters dalyvavimą darbo rinkoje.
3. Kokias išvadas galite daryti remdamiesi šiuo modeliu?
4. Heckmano modelis atlyginimui w_i ir dalyvavimui darbo rinkoje h_i aprašomas šiomis lygtimis:

$$w_i^* = \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_{1i} \quad (1)$$

$$h_i^* = \mathbf{z}'_i \boldsymbol{\gamma} + \varepsilon_{2i} \quad (2)$$

$$w_i = w_i^*, h_i = 1, \text{ jei } h_i^* > 0 \quad (3)$$

$$w_i \text{ nestebimas}, h_i = 0, \text{ jei } h_i^* \leq 0, \quad (4)$$

čia

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & 1 \end{pmatrix} \right)$$

Taigi turime modelį su dviem endogeniniais kintamaisiais w_i ir h_i . Koeficientus $\boldsymbol{\beta}$ ir $\boldsymbol{\gamma}$ galima įvertinti dviem būdais: didžiausio tikėtinumo ir „heckit“. „Heckit“ metodas susideda iš dviejų etapų. Pirmame etape yra įvertinama lygtis (2), naudojant probit regresiją, antrame mažiausių kvadratų metodu įvertinama regresija:

$$w_i = \mathbf{x}'_i \boldsymbol{\beta} + \sigma_{12} \lambda_i + \eta_i,$$

čia

$$\lambda_i = \frac{\varphi(\mathbf{z}'_i \boldsymbol{\gamma})}{\Phi(\mathbf{z}'_i \boldsymbol{\gamma})}.$$

Šio modelio paklaidos yra lygios $\eta_i = \varepsilon_{1i} - \mathbf{E}(\varepsilon_{1i} | x_i, h_i = 1)$.

Parašykite funkciją, kuri atliktų heckit procedūrą, ir įvertinkite jūsų pasirinktą modelį. Kodėl negalima naudoti koeficientų paklaidų iš įvertintos tiesinės regresijos?

5. Didžiausio tikėtinumo funkcija Heckman modeliui atrodo taip:

$$\log L(\boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_1^2, \sigma_{12}) = \sum_{h_i=0} \log P(h_i = 0) + \sum_{h_i=1} [\log f(w_i) + \log P(h_i = 1 | w_i)],$$

čia

$$\begin{aligned} P(h_i = 0) &= 1 - \Phi(\mathbf{z}'_i \boldsymbol{\gamma}) \\ P(h_i = 1 | w_i) &= \Phi \left(\frac{\mathbf{z}'_i \boldsymbol{\gamma} + (\sigma_{12}/\sigma_1^2)(w_i - \mathbf{x}'_i \boldsymbol{\beta})}{\sqrt{1 - \sigma_{12}^2/\sigma_1^2}} \right) \\ f(w_i) &= \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp \left\{ -\frac{1}{2}(w_i - \mathbf{x}'_i \boldsymbol{\beta})^2 / \sigma_1^2 \right\} \end{aligned}$$

Parašykite funkciją įvertinančią didžiausio tikėtinumo funkcijos parametrus.

6. Įvertinkite modelį, kurį vertinote su heckit, didžiausio tikėtinumo metodu. Kiek skiriasi parametrai? Įvertinkite parametru standartines paklaidas ir patikrinkite, kurie parametrai yra reikšmingi. (Naudokitės didžiausio tikėtinumo įverčių savybėmis). Įvertinkite savo modelio gerumą, bei adekvatumą.