

OCENJEVANJE PARAMETROV V PROBIT MODELU

Pri PROBIT modelu privzemamo, da imamo za vsako enoto v populaciji samo odziv 0 ali 1, verjetnost za odziv 1 (ali 0) pa je še odvisna od neodvisnih spremenljivk X_1, X_2, \dots, X_m in sicer kot

$$P(Y = 1|X_1, X_2, \dots, X_m) = \Phi(\alpha + \beta_1 X_1 + \dots + \beta_m X_m),$$

kjer je $\Phi(x)$ porazdelitvena funkcija standardne normalne porazdelitve. Ko izberemo iz populacije vzorec, privzemamo, da je za i -to enoto

$$P(Y_i = 1) = \Phi(\alpha + \beta_1 x_{i1} + \dots + \beta_m x_{im}),$$

kjer je Y_i odziv i -te enote x_{i1}, \dots, x_{im} pa so neodvisne spremenljivke za i -to enoto, odzivi enot pa so med sabo neodvisne slučajne spremenljivke. Vemo, da je logaritem funkcije verjetja enak

$$\begin{aligned} \ell(\mathbf{y}, \mathbf{X}, \alpha, \beta) &= \sum_{i=1}^n y_i \log(\Phi(\alpha + \beta_1 x_{i1} + \dots + \beta_m x_{im})) \\ &+ \sum_{i=1}^n (1 - y_i) \log(1 - \Phi(\alpha + \beta_1 x_{i1} + \dots + \beta_m x_{im})). \end{aligned}$$

Velikost vzorca je tukaj n .

Priložen je izpis iz programa LINDEP za podatke, ki so na razpolago na moji domači strani. Na podlagi izpisa in podatkov naredite naslednje:

- Preverite, da sta parametra α in β ocenjena po metodi največjega verjetja.
- Izračunajte

$$\begin{aligned} f_{11} &= E\left(-\frac{\partial^2 \ell(\mathbf{Y}, \mathbf{X}, \alpha, \beta)}{\partial \alpha^2}\right) \\ f_{12} &= E\left(-\frac{\partial^2 \ell(\mathbf{Y}, \mathbf{X}, \alpha, \beta)}{\partial \alpha \partial \beta}\right) \quad \text{in} \\ f_{22} &= E\left(-\frac{\partial^2 \ell(\mathbf{Y}, \mathbf{X}, \alpha, \beta)}{\partial \beta^2}\right). \end{aligned}$$

Tukaj je $\ell(\mathbf{Y}, \mathbf{X}, \alpha, \beta)$ celotna funkcija verjetja. Prepričajte se, da sta standardni napaki za oceni parametrov na izpisu dobljeni iz varianc, ki jih preberemo iz inverzne matrike informacije

$$\mathbf{F}^{-1} = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix}^{-1}$$

Poskusite pojasniti, zakaj dobimo iz zgornje inverzne matrike aproksimaciji za varianci cenilk obeh parametrov. Privzeti morate nekoliko splošnejši centralni limitni izrek kot tisti, ki smo ga obravnavali na predavanjih.

- c. Na izpisu je tudi $\chi^2(1)$ statistika. Formulirajte domnevo, ki jo preverjamo s to statistiko in preverite, da je dobljena po metodi kvocienta verjetij. Pojasnite, zakaj ima eno prostostno stopnjo.
- d. Zakaj ni nobenega preizkusa domnev za parameter α ?
- e. Kako bi ustrezno formulirali doslednost cenilk v tem primeru? Je po vašem mnenju cenilka po metodi največjega verjetja dosledna?

```
MODEL COMMAND: PROBIT;LHS==Y;RHS==ONE,X$
```

```
Method==NEWTON; Maximum iterations == 25
Convergence criteria: Gradient == .1000000E-03
Function == .1000000E-05
Parameters== .1000000E-04
Starting values: .7212 -.2109
```

```
====> NEWTON Iterations
```

Iteration	1	Function	212.1327
Param	.721	-.211	
Gradnt	19.5	87.1	
Iteration	2	Function	189.2619
Param	.686	-.694	

Gradnt -4.33 9.75

Iteration 3 Function 188.7442
Param .728 -.778
Gradnt -.276 .416

Iteration 4 Function 188.7430
Param .731 -.782
Gradnt -.760E-03 .102E-02

Iteration 5 Function 188.7430
Param .731 -.782
Gradnt -.510E-08 .644E-08

** Gradient has converged.
** Function has converged.
** B-vector has converged.

Maximum Likelihood Estimates

Log-Likelihood..... -188.74
Restricted (Slopes==0) Log-L. -233.30
Chi-Squared (1)..... 89.121
Significance Level..... .32173E-13

Variable Coefficient Std. Error T-ratio Prob|t|F2x Mean of X
Std.D.of X

ONE	.730790	.775747E-01	9.420	.00000	1.0000
	.00000				
X	-.782305	.925799E-01	-8.450	.00000	-.41779E-01
	.94192				