

## Modele ograniczonej zmiennej zależnej. Funkcja probitowa i logitowa.

Zmienną objaśnianą jest zmienna jakościowa wyrażona przez zmienną zero-jedynkową. Zmienna objaśniana może przyjmować dwie wartości 0 lub 1.

Wartość oczekiwana tej zmiennej może być interpretowana jako prawdopodobieństwo warunkowe danego zdarzenia, przy ustalonych wartościach zmiennych objaśniających.

Jeżeli zmienna objaśniana jest zmienną jakościową, to nie można stosować klasycznych modeli ekonometrycznych, ponieważ wartości zmiennych (wartości prawdopodobieństwa) mogą przyjmować wartości jedynie z przedziału (0,1).

Dokonuje się odpowiednich monotonicznych przekształceń przedziału prawdopodobieństwa (0,1) na przedział  $(-\infty, +\infty)$ . Najpopularniejszymi są przekształcenia probitowe i logitowe.

Ponieważ zmienna objaśniana przyjmuje tylko dwie wartości, konieczne jest wprowadzenie pewnej liczby kategorii zmiennej objaśniającej (konstrukcja szeregu rozdzielczego), tak aby można było mierzyć częstość wystąpienia wariantu zmiennej objaśnianej w każdej z tych kategorii.

### Transformacja probitowa

W miejsce zmiennej objaśnianej podstawia się wartości probitów:

$$Pr = F^{-1}(p) + 5,$$

gdzie  $F^{-1}$  – funkcja odwrotna do dystrybuanty standaryzowanego rozkładu normalnego, liczba 5 jest dodawana w celu uniknięcia wartości ujemnych.

Za  $p$  przyjmuje się częstości oszacowane na podstawie próby.

Parametry modelu probitowego:

$$Pr_i = \alpha_0 + \alpha_1 X_{1i} + \varepsilon_i$$

szacuje się za pomocą UMNK, a przy małej liczbie obserwacji za pomocą metody największej wiarygodności (wariancja składnika losowego nie jest stała).

Elementy macierzy diagonalnej  $V$  oblicza się następująco (przy UMNK):

$$v_i = p_i (1-p_i) / n_i \varphi^2(p_i),$$

gdzie  $\varphi(p_i)$  jest funkcją gęstości standaryzowanego rozkładu normalnego.

Aby otrzymać na podstawie oszacowanego modelu prawdopodobieństwa  $p_i$  należy obliczyć wartość dystrybuanty  $F$ :

$$p_i = F(a_0 + a_1 X_{1i} - 5).$$

### Transformacja logitowa

W miejsce zmiennej objaśnianej podstawia się wartości logitów:

$$L = \ln \frac{p}{1-p}.$$

Parametry modelu logitowego:

$$L_i = \alpha_0 + \alpha_1 X_{1i} + \varepsilon_i$$

szacuje się za pomocą UMNK, a przy małej liczbie obserwacji za pomocą metody największej wiarygodności (wariancja składnika losowego nie jest stała).

Elementy macierzy diagonalnej  $V$  oblicza się następująco (przy UMNK):

$$v_i = 1 / n_i p_i (1-p_i).$$

Aby otrzymać na podstawie oszacowanego modelu prawdopodobieństwa  $p_i$  należy sprowadzić go do postaci logistycznej:

$$p_i = \frac{1}{1 + \exp[-(a_0 + a_1 x_{1i})]}.$$

Zadanie

Pewien urząd pracy zlecił sporządzenie prognozy szans znalezienia pracy przez osobę bezrobotną w ciągu roku od momentu zarejestrowania się w tym urzędzie.

i	$n_i$	Wiek w latach	Średni staż w latach	$m_i$
1	1000	18-22	1	100
2	1500	22-26	3	160
3	900	26-30	5	110
4	800	30-34	7	130
5	1000	34-38	9	180
6	800	38-42	20	200
7	400	42-46	20	110
8	200	46-50	25	60
9	100	50-54	28	33
10	40	54-58	27	13

$n_i$  – liczba bezrobotnych,

$m_i$  – liczba bezrobotnych w danej grupie, którzy w ciągu roku znaleźli pracę.

Polecana literatura:

Zeliaś A., Pawełek B., Wanat S., Prognozowanie ekonomiczne. Teoria, przykłady, zadania.