Uniwersytet Ekonomiczny w Katowicach

Studia Ekonomiczne. Zeszyty Naukowe
Uniwersytetu Ekonomicznego w Katowicach
ISSN 2083-8611 Nr 324 · 2017

Jerzy W. Wiśniewski

Uniwersytet Mikołaja Kopernika
Wydział Nauk Ekonomicznych i Zarządzania
Katedra Ekonometrii i Statystyki
jerzy.wisniewski@umk.pl

LINIOWA FUNKCJA PRAWDOPODOBIEŃSTWA

Streszczenie: Ważnym narzędziem analitycznym w ekonometrii, służącym m.in. do badania asocjacji zmiennych zerojedynkowych, jest liniowa funkcja prawdopodobieństwa, zwana też modelem Goldbergera[1]. Jego specyfiką jest zerojedynkowa zmienna objaśniana, powodująca, że teoretyczne wartości modelu empirycznego są szacunkami prawdopodobieństwa wystąpienia wariantu sygnowanego liczbą 1. Zmienne objaśniające w modelu mogą być zarówno ciągłe, jak i dyskretne. Model Goldbergera jest ważnym instrumentem pomiaru uwarunkowań przyczynowych, głównie zmiennych jakościowych, ale również tzw. zmiennych ilościowych. Wymaga on jednak specyficznego podejścia, przede wszystkim do estymacji jego parametrów.

Słowa kluczowe: zmienna zerojedynkowa, asocjacja, prawdopodobieństwo, model ekonometryczny.

JEL Classification: C01, C25, C51.

Wprowadzenie

Model ekonometryczny, noszący nazwę liniowej funkcji prawdopodobieństwa, nie znalazł należytej akceptacji w środowisku ekonomistów. Charakteryzuje się tym, że zmienna objaśniana równania regresji jest zerojedynkową. Może być zatem bardzo dobrym narzędziem analizy asocjacji cech jakościowych[2], lepszym – w sensie walorów informacyjnych – od rozmaitych współczynników asocjacji.

---

1 Por. [Goldberger 1972, s. 319-321; Wiśniewski, 1990; Wiśniewski, 2012].

2 O pomiarze cech jakościowych traktują prace: [Churgin, 1985; Stevens, 1946; Wiśniewski, 2013b].

Celem niniejszej pracy jest wskazanie walorów decyzyjnych liniowej funkcji prawdopodobieństwa, zwanej tez modelem Goldbergera. Ważnym zadaniem badawczym jest też prezentacja konieczności stosowania wyspecjalizowanych narzędzi w procesie estymacji jego parametrów. Owe szczególne procedury będą konieczne dla uzyskania odpowiednio precyzyjnych wyników w postaci empirycznego modelu Goldbergera.

1. Istota modelu Goldbergera

Zmienna objaśniana modelu ekonometrycznego winna charakteryzować się relatywnie dużym obszarem zmienności. Nie powinna też być ograniczona. Oznacza to, że nie powinna posiadać ani dolnej, ani górnej granicy. Tymczasem niekiedy pojawiają się zmienne, pełniące w modelu rolę objaśnianych, o obserwacjach $y_i^{(o)}$, które posiadają nawet dwustronne ograniczenia. Ich specyfiką jest posiadanie dolnej i górnej granicy, czyli:

$$y_{min} \leq y_i^{(o)} \leq y_{max}, \tag{1}$$

gdzie y_{min} oznacza najniższą możliwą wartość obserwacji na rozważanej zmiennej, natomiast y_{max} jest najwyższą możliwą wartością obserwacji na tej zmiennej objaśnianej.

Załóżmy, że mechanizm zmienności zmiennej ograniczonej $y_i^{(o)}$ będzie opisany za pomocą modelu liniowego:

$$y_i^{(o)} = \alpha_0 + \alpha_1 x_i + \eta_i. \tag{2}$$

Rysunek 1 prezentuje liniowy model ekonometryczny dla ograniczonej zmiennej objaśnianej. Zauważa się konsekwencje ewentualnej ekstrapolacji poza obszar obserwacji statystycznych. Próba takiej ekstrapolacji może prowadzić do tego, że wartości ekstrapolant będą mieściły się poza obszarem zmienności zmiennej ograniczonej, co jest sprzeczne z logiką. Na przykład, zmienną ograniczoną może być wskaźnik struktury, spełniający nierówność $0 \leq y_i^{(o)} \leq 100.$ Próba ekstrapolacji zmiennej w postaci wskaźnika struktury może doprowadzić do tego, że ekstrapolanty osiągną wartości mniejsze od 0% albo większe od 100%, co jest sprzeczne z istotą tego wskaźnika.

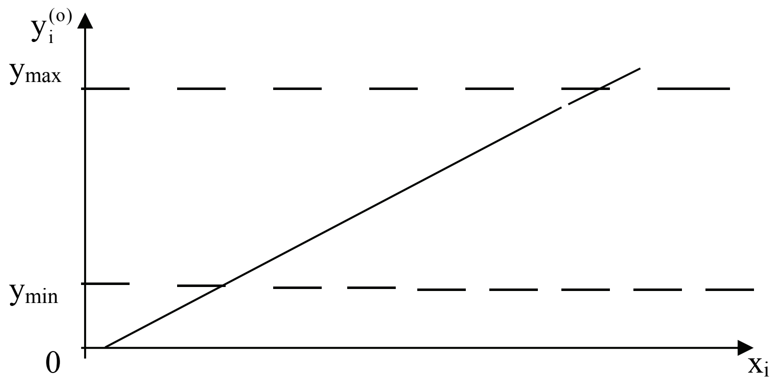
Szczególnym przypadkiem modelu z ograniczoną zmienną objaśnianą jest liniowa funkcja prawdopodobieństwa, zwana też modelem Goldbergera[3]. Model ten można zapisać następująco:

$$y_i = \alpha_0 + \alpha_1 x_{i1} + ... + \alpha_j x_{ij} + ... + \alpha_k x_{ik} + \eta_i, \quad (3)$$

gdzie:
y_i – zmienna zerojedynkowa, zdefiniowana następująco:

$$y_i = \begin{cases} 1, & \text{przy warunkach } V, \\ 0, & \text{w przypadku przeciwnym}, \end{cases}$$

$x_{i1}, \ldots, x_{ij}, \ldots, x_{ik}$ – obserwacje na zmiennych objaśniających,
η_i – składnik losowy równania,
$\alpha_0, \alpha_1, \ldots, \alpha_j, \ldots, \alpha_k$ – parametry strukturalne modelu,
i – numer obserwacji statystycznej (i = 1, …, n).



Rys. 1. Liniowy model ograniczonej zmiennej objaśnianej

Zmienne objaśniające w modelu (3) mogą być zarówno ciągłe, jak i dyskretne. W tym zbiorze zmiennych mogą pojawiać się również zmienne zerojedynkowe, będące szczególnym przypadkiem zmiennych skokowych.

Skonstruowanie empirycznej liniowej funkcji prawdopodobieństwa pozwala na ustalenie częstości pojawiania się warunków V, przy określonych konfiguracjach wartości zmiennych objaśniających. W empirycznym modelu Goldbergera możliwe jest też wskazanie statystycznie istotnych zmiennych objaśniających, które mają wpływ na zaistnienie warunków V.

[3] Jedną z pierwszych monografii w Polsce traktującą o modelu Goldbergera jest: [Wiśniewski, 1986]. Empiryczne rezultaty zastosowań tego modelu znaleźć można w pracach: [Wiśniewski, 2009, 2013a].

2. Charakterystyka procedury estymacyjnej

Estymacja parametrów modelu (2) za pomocą klasycznej metody najmniejszych kwadratów (KMNK) skutkuje ograniczoną precyzją szacunków parametrów[4]. Rezultatem zastosowania KMNK jest przypadek niejednorodności wariancji składnika losowego.

Konieczna jest w tym przypadku procedura dwustopniowa. W pierwszym kroku należy zastosować KMNK do oszacowania parametrów modelu z zerojedynkową zmienną endogeniczną. Po obliczeniu teoretycznych wartości z równania empirycznego typu (3) można wyznaczyć wagi dla każdej obserwacji, obliczane następująco:

$$w_i = \hat{y}_i(1-\hat{y}_i), \qquad (i=1, \quad ..., \quad n) \tag{4}$$

W rezultacie można skonstruować empiryczną macierz $\hat{\Omega}$ o postaci:

$$\hat{\Omega} = \begin{bmatrix} w_1 & ... & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & ... & w_i & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & ... & 0 & ... & w_n \end{bmatrix}. \tag{5}$$

W praktyce mogą pojawić się ujemne wartości wag w_i. Dlatego też lepszym wariantem będzie wykorzystanie modułów wag obliczonych według wzoru (4). Macierz $\hat{\Omega}$ przyjmie więc następującą postać:

$$\hat{\Omega} = \begin{bmatrix} |w_1| & ... & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & ... & |w_i| & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & ... & 0 & ... & |w_n| \end{bmatrix}. \tag{6}$$

---

4 W niniejszej pracy zastosowano termin „precyzja szacunków", zamiast stosowanego w literaturze pojęcia „efektywności szacunków". Wynika to z istnienia w statystyce pojęcia „precyzji τ", obliczanej jako: $\tau = \frac{1}{\sigma}$, gdzie σ jest odchyleniem standardowym. Spadek wartości odchylenia standardowego oznacza poprawę efektywności estymacji. Ze spadkiem wielkości σ następuje wzrost precyzji τ, wskazujący na poprawę precyzji estymatora.

Estymator Aitkena dla przypadku objaśnianej zmiennej zerojedynkowej będzie miał zatem następującą postać:

$$\overset{*}{\alpha} = (X^T \hat{\Omega}^{-1} X)^{-1} X^T \hat{\Omega}^{-1} Y \tag{7}$$

Macierz $\hat{\Omega}^{-1}$ będzie miała następującą strukturę:

$$\hat{\Omega}^{-1} = \begin{bmatrix} \left|\frac{1}{w_1}\right| & \ldots & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \left|\frac{1}{w_i}\right| & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \ldots & \left|\frac{1}{w_n}\right| \end{bmatrix}$$

albo

$$\hat{\Omega}^{-1} = \begin{bmatrix} \left|\frac{1}{\hat{y}_1(1-\hat{y}_1)}\right| & \ldots & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \left|\frac{1}{\hat{y}_i(1-\hat{y}_i)}\right| & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \ldots & \left|\frac{1}{\hat{y}_n(1-\hat{y}_n)}\right| \end{bmatrix} \tag{8}$$

Estymator (7) daje bardziej efektywne (precyzyjne) szacunki parametrów modelu z zerojedynkową zmienną objaśnianą, w porównaniu z estymatorem KMNK.

3. Empiryczny model Goldbergera

Poniżej zaprezentowano przykład konstrukcji empirycznego modelu Goldbergera. Zagadnienie decyzyjne polega na ustaleniu cech osobistych handlowców w przedsiębiorstwie ZET, które mają istotny statystycznie wpływ na generowanie przez nich wierzytelności przeterminowanych[5]. Informacje statystyczne o efektywności pracy handlowców oraz ich cechach osobistych zawiera tab. 1. Liniowa funkcja prawdopodobieństwa opisywała będzie skuteczność windykacji wierzytelności, w zależności od rozmaitych cech osobistych handlowców.

Tabela 1. Efektywność pracy handlowców i ich cechy osobiste w przedsiębiorstwie ZET

Nr	y1i	y2i	xi1	xi2	xi3	xi4	xi5	xi6	xi7
1	0	545	1	0	3	0	0	24	0
2	0	550	0	0	3	0	0	25	0
3	0	563	0	0	2	0	0	23	0
4	0	569	0	0	4	0	2	25	0
5	0	570	0	0	3	0	1	27	1
6	0	581	1	1	5	0	1	26	1
7	1	583	0	0	5	0	1	28	1
8	0	588	0	0	5	1	3	27	0
9	0	591	1	1	3	0	0	29	0
10	0	594	1	0	4	0	0	30	0
11	0	595	1	0	7	0	1	30	0
12	0	597	0	0	6	0	1	29	0
13	1	600	0	1	5	0	1	31	1
14	0	604	0	0	8	0	1	32	0
15	0	604	0	0	6	0	2	33	0
16	0	605	1	0	8	1	1	34	0
17	0	606	0	0	8	1	2	33	0
18	1	616	0	1	7	0	1	35	1
19	0	619	0	0	9	0	1	36	1
20	0	620	1	0	8	0	2	36	0
21	0	624	1	0	9	0	2	34	1
22	0	625	0	0	9	0	3	36	1
23	1	638	0	1	7	0	1	37	1
24	0	641	0	0	8	1	2	38	0
25	0	644	0	0	9	1	2	39	1
26	0	663	1	0	10	0	3	40	1
27	0	688	1	1	10	1	3	40	0
28	0	725	0	1	11	1	4	41	0
29	1	753	0	0	12	0	4	40	0
30	0	788	0	1	11	1	2	39	0
31	0	801	0	0	9	0	1	35	0
32	1	803	0	0	8	0	3	37	0
33	1	821	0	1	12	0	3	42	1
34	0	843	1	1	17	1	5	47	1
35	0	866	1	0	8	0	1	31	0

Źródło: dane przedsiębiorstwa ZET analogiczne do rzeczywistych.

Pomiaru skuteczności windykacji wierzytelności dokonano za pomocą zmiennej zerojedynkowej, zdefiniowanej następująco[6]:

---

[5] Przykład jest analogiczny do zamieszczonego w monografii: [Wiśniewski, 2016, podrozdział 6.4].

[6] Handlowiec winien uzyskiwać wysokie przychody ze sprzedaży oraz zabiegać o skuteczną windykację wierzytelności w obsługiwanej przez niego sieci sprzedaży.

y_{1i} – zmienna zerojedynkowa przyjmująca wartość 1, gdy w sieci sprzedaży i-tego handlowca powstały wierzytelności przeterminowane[7] oraz zero w przypadku przeciwnym,

y_{2i} – przychód ze sprzedaży netto uzyskany rocznie przez i-tego handlowca (tys. zł),

x_{i1} – zmienna zerojedynkowa, reprezentująca płeć handlowca, przyjmująca wartość 1 dla kobiet i 0 dla mężczyzn,

x_{i2} – zmienna zerojedynkowa, informująca o fakcie uprawiania sportu wyczynowego przez handlowca, przyjmująca wartość 1, gdy uprawiał on sport wyczynowo oraz 0 w przypadku przeciwnym,

x_{i3} – staż pracy w zawodzie handlowca, wyrażony liczbą przepracowanych lat,

x_{i4} – zmienna zerojedynkowa, informująca o posiadaniu wykształcenia ekonomicznego, przyjmująca wartość 1, gdy handlowiec posiada wykształcenie ekonomiczne oraz 0, gdy takiego wykształcenia nie posiada,

x_{i5} – liczba osób na utrzymaniu handlowca,

x_{i6} – wiek handlowca, wyrażony liczbą ukończonych lat życia,

x_{i7} – zmienna zerojedynkowa, informująca o posiadaniu wykształcenia wyższego, przyjmująca wartość 1, gdy handlowiec posiada wykształcenie wyższe oraz 0, gdy takiego wykształcenia nie posiada.

Rozważono hipotetyczny model rekurencyjny[8], który zawiera liniową funkcję prawdopodobieństwa, opisującą y_{1i}:

$$\begin{aligned} y_{1i} &= \alpha_{10} + \beta_{12} y_{2i} + \alpha_{11} x_{i1} + \alpha_{12} x_{i2} + \alpha_{13} x_{i3} + \\ &+ \alpha_{14} x_{i4} + \alpha_{15} x_{i5} + \alpha_{16} x_{i6} + \alpha_{17} x_{i7} + \eta_{1i}, \end{aligned} \tag{9}$$

$$\begin{aligned} y_{2i} &= \alpha_{20} + \alpha_{21} x_{i1} + \alpha_{22} x_{i2} + \alpha_{23} x_{i3} + \\ &+ \alpha_{24} x_{i4} + \alpha_{25} x_{i5} + \alpha_{26} x_{i6} + \alpha_{27} x_{i7} + \eta_{2i}, \end{aligned} \tag{10}$$

gdzie[9]:

α_{1i}, α_{2i} – parametry strukturalne przy zmiennych z góry ustalonych modelu (i=0,1,…,7),

β_{12} – parametr strukturalny przy zmiennej łącznie współzależnej modelu,

η_{1i}, η_{2i} – składniki losowe równań modelu.

---

[7] Chodzi o należności przeterminowane ponad ustaloną w przedsiębiorstwie normę.

[8] Pojawienie się w równaniu (9) zmiennej objaśniającej, wyrażającej wartość przychodów ze sprzedaży, uzyskiwanych przez handlowca, wynika z wątpliwości, czy system motywowania jest poprawnie skonstruowany. Poprawny system motywacyjny nie powinien zachęcać do tworzenia wierzytelności przeterminowanych w części sieci, którą obsługuje handlowiec. W dalszej części niniejszej pracy zajęto się estymacją parametrów liniowej funkcji prawdopodobieństwa, czyli równania (9), rezygnując z rozważań nad równaniem (10).

[9] Por. [Wiśniewski, 2016, s. 187].

Parametry równania (8) oszacowano za pomocą klasycznej metody najmniejszych kwadratów (KMNK), z wykorzystaniem pakietu GRETL. Po eliminacji zmiennych objaśniających, które okazały się statystycznie nieistotne, powstał model empiryczny, którego charakterystyki zawiera tab. 2. Reszty równania empirycznego:

$$\hat{y}_{1i} = -0{,}3189 - 0{,}3318x_{i1} + 0{,}305x_{i2} - 0{,}4494x_{i4} + 0{,}0198x_{i6} \qquad (11)$$

wykorzystane zostały do konstrukcji wag typu (4).

Tabela 2. Estymacja KMNK[10], wykorzystane obserwacje 1-35

Zmienna zależna (Y): y_{1i}

Wyszczególnienie	Współczynnik	t-Studenta	Wartość p	Ważność
const	-0,318944	-0,9155	0,3672	
x_{i1}	-0,331834	-2,9057	0,0068	***
x_{i2}	0,304972	2,3646	0,0247	**
x_{i4}	-0,449381	-3,2852	0,0026	***
x_{i6}	0,0197945	1,8179	0,0791	*

Średn. aryt. zm. zależnej	0,200000	Odch. stand. zm. zależnej	0,405840
Suma kwadratów reszt	3,064098	Błąd standardowy reszt	0,319588
Wsp. determ. R-kwadrat	0,452840	Skorygowany R-kwadrat	0,379885
F(4, 30)	6,207134	Wartość p dla testu F	0,000919
Logarytm wiarygodności	-7,039938	Kryt. inform. Akaike'a	24,07988
Kryt. bayes. Schwarza	31,85662	Kryt. Hannana-Quinna	26,76441

Tabela 3. Estymacja WLS, wykorzystane obserwacje 1-35

Zmienna zależna (Y): y_{1i} Zmienna jako waga: wagi

Wyszczególnienie	Współczynnik	Błąd stand.	t-Studenta	Wartość p	Ważność
const	-0,0580519	0,236949	-0,2450	0,8081	
x_{i1}	-0,304019	0,0680465	-4,4678	0,0001	***
x_{i2}	0,295007	0,0764104	3,8608	0,0006	***
x_{i4}	-0,404966	0,0865817	-4,6773	< 0,0001	***
x_{i6}	0,0117954	0,00746274	1,5806	0,1245	

[10] Trzy gwiazdki w kolumnie ważność oznaczają statystyczną istotność na poziomie istotności poniżej p = 0,01; dwie gwiazdki oznaczają istotność zmiennej na poziomie istotności poniżej p = 0,05, natomiast jedna gwiazdka oznacza istotność na poziomie istotności poniżej p = 0,10. Brak gwiazdki oznacza brak statystycznej istotności danej zmiennej objaśniającej.

cd. tabeli 3

Podstawowe statystki dla ważonych danych:

Suma kwadratów reszt	15,62920		Błąd standardowy reszt	0,721785
Wsp. determ. R-kwadrat	0,524961		Skorygowany R-kwadrat	0,461623
F(4, 30)	8,288193		Wartość p dla testu F	0,000126
Logarytm wiarygodności	-35,55423		Kryt. inform. Akaike'a	81,10845
Kryt. bayes. Schwarza	88,88519		Kryt. Hannana-Quinna	83,79298

Podstawowe statystyki dla oryginalnych danych:

Średn. aryt. zm. zależnej	0,200000		Odch. stand. zm. zależnej	0,405840
Suma kwadratów reszt	3,140178		Błąd standardowy reszt	0,323531

Porównanie wyników z tab. 3 i 2 wskazuje na wzrost wartości współczynnika R^2 w rezultacie zastosowania uogólnionej metody najmniejszych kwadratów (do poziomu 0,525) w porównaniu z KMNK (0,4528). Ponadto wzrosły wartości statystyk t-Studenta, związane ze zmiennymi x_{i1}, x_{i2} oraz x_{i4}. Po eliminacji „słabej" zmiennej objaśniającej x_{i6} otrzymujemy wynik, zamieszczony w tab. 4.

Tabela 4. Estymacja WLS, wykorzystane obserwacje 1-35

Zmienna zależna (Y): y_{1i}. Zmienna jako waga: wagi

Wyszczególnienie	Współczynnik	Błąd stand.	t-Studenta	Wartość p	Ważność
const	0,30457	0,0606509	5,0217	< 0,0001	***
x_{i1}	-0,322538	0,0686309	−4,6996	< 0,0001	***
x_{i2}	0,3322	0,0744326	4,4631	< 0,0001	***
x_{i4}	-0,312774	0,0655137	−4,7742	< 0,0001	***

Podstawowe statystki dla ważonych danych:

Suma kwadratów reszt	16,93070		Błąd standardowy reszt	0,739021
Wsp. determ. R-kwadrat	0,485403		Skorygowany R-kwadrat	0,435604
F(3, 31)	9,747119		Wartość p dla testu F	0,000110
Logarytm wiarygodności	-36,95401		Kryt. inform. Akaike'a	81,90801
Kryt. bayes. Schwarza	88,12941		Kryt. Hannana-Quinna	84,05564

Podstawowe statystyki dla oryginalnych danych:

Średn. aryt. zm. zależnej	0,200000		Odch. stand. zm. zależnej	0,405840
Suma kwadratów reszt	3,425596		Błąd standardowy reszt	0,332420

4. Rozwiązanie uzupełniające dla modelu Goldbergera

Posiadanie szacunków prawdopodobieństwa (częstości) tworzenia wierzytelności przeterminowanych $(\hat{y}_{1i})$ pozwala na zastosowanie innego rozwiązania modelowego dla przygotowania narzędzia decyzyjnego. Owe częstości wykorzystane zostaną do konstrukcji równania empirycznego, w którym zmienną objaśnianą będzie logitowa transformacja zmiennej ograniczonej, jaką jest $\hat{y}_{1i}$. Przekształcenie tej zmiennej ograniczonej odbędzie się w dwóch krokach. Najpierw wykonana zostanie transformacja podstawowa:

$$y_{1i}^{(p)} = \frac{\hat{y}_{1i} - \hat{y}_{1i\,min}}{\hat{y}_{1i\,max} - \hat{y}_{1i}} \tag{12}$$

która powoduje, że zmienna dwustronnie ograniczona staje się ograniczoną jednostronnie, z minimum wynoszącym zero. Kolejna transformacja logitowa:

$$y_{1i}^{(l)} = \ln y_{1i}^{(p)} = \ln \frac{\hat{y}_{1i} - \hat{y}_{1i\,min}}{\hat{y}_{1i\,max} - \hat{y}_{1i}} \tag{13}$$

przekształca częstości w zmienną nieograniczoną. Rozważono zatem równanie:

$$\begin{aligned} y_{1i}^{(l)} &= \alpha_{10} + \beta_{12} y_{2i} + \alpha_{11} x_{i1} + \alpha_{12} x_{i2} + \alpha_{13} x_{i3} + \\ &+ \alpha_{14} x_{i4} + \alpha_{15} x_{i5} + \alpha_{16} x_{i6} + \alpha_{17} x_{i7} + \eta_{1i} \end{aligned} \tag{14}$$

którego parametry oszacowane zostaną za pomocą KMNK. Rezultaty estymacji zamieszczone zostały w tab. 5. Dostrzega się kolejną poprawę wartości współczynnika R^2, którego wartość wzrosła do poziomu 0,585. Ponadto poprawiła się precyzja szacunków parametrów strukturalnych równania.

Tabela 5. Estymacja KMNK, wykorzystane obserwacje 1-35

Zmienna zależna (Y): $y_{1i}^{(l)}$

Wyszczególnienie	Współczynnik	Błąd stand.	t-Studenta	Wartość p	Ważność
const	-1,02934	0,544184	-1,8915	0,0682	*
x_{i1}	-0,657637	0,178386	-3,6866	0,0009	***
x_{i2}	0,734861	0,201462	3,6476	0,0010	***
x_{i4}	-0,870824	0,213664	-4,0757	0,0003	***
x_{i6}	0,0325102	0,0170084	1,9114	0,0655	*

cd. tabeli 5

Średn. aryt. zm. zależnej	-0,182939		Odch. stand. zm. zależnej	0,727892
Suma kwadratów reszt	7,476064		Błąd standardowy reszt	0,499201
Wsp. determ. R-kwadrat	0,584988		Skorygowany R-kwadrat	0,529653
F(4, 30)	10,57178		Wartość p dla testu F	0,000018
Logarytm wiarygodności	-22,64912		Kryt. inform. Akaike'a	55,29824
Kryt. bayes. Schwarza	63,07498		Kryt. Hannana-Quinna	57,98277

Empiryczna funkcja prawdopodobieństwa winna dostarczyć informacji, które będą przydatne w podejmowaniu decyzji. W rozpatrywanym przypadku kolejne rozwiązania modelowe udzieliły odpowiedzi na pytanie o istotne statystycznie zmienne, wpływające na skuteczność windykacji handlowca. Cztery cechy osobiste okazały się znaczące w działaniach windykacyjnych handlowca, z których trzy można uznać za decydujące. Rozwiązanie zawarte w tab. 5 pozwoliło uzyskać ważną informację na temat istotnie statystycznym oddziaływaniu zmiennej x_{i6} na prawdopodobieństwo generowania przez handlowca wierzytelności przeterminowanych. Poprawiła się zatem wartość poznawcza wyników empirycznych. Wykorzystanie tych informacji w konkretnym przedsiębiorstwie, gdy pojawi się potrzeba zatrudnienia kolejnego handlowca, pozwoli na ograniczenie ryzyka błędnej decyzji kadrowej.

Podsumowanie

Niniejszy artykuł pozwala na wysunięcie następujących wniosków:

1. Liniowa funkcja prawdopodobieństwa jest przydatnym narzędziem analitycznym ekonomisty.
2. Umożliwia wskazanie istotnych statystycznie zmiennych wpływających na pojawienie się warunków zdefiniowanych za pomocą zmiennej zerojedynkowej w wariancie przyjmującym wartość jeden.
3. Może być wykorzystana w analizie tzw. cech jakościowych, jak też we wskazywaniu uwarunkowań pojawiania się tzw. obserwacji nietypowych w szeregach statystycznych.
4. Umiejętne wykorzystanie empirycznego modelu Goldbergera pozwoli na podejmowanie decyzji, charakteryzujących się zmniejszeniem ryzyka ich negatywnych następstw.

Literatura

Churgin J. (1985), *Jak policzyć niepoliczalne*, Wiedza Powszechna, seria OMEGA, Warszawa.

Goldberger A.S. (1972), *Teoria ekonometrii*, PWE, Warszawa.

Stevens S.S. (1946), *On the Theory of Scales Measurement*, "Science", Vol. 103, No. 2684.

Wiśniewski J.W. (1986), *Ekonometryczne badanie zjawisk jakościowych. Studium metodologiczne*, UMK, Toruń.

Wiśniewski J.W. (1990), *Dynamiczne modelowanie ekonometryczne ograniczonej zmiennej zależnej*, „Przegląd Statystyczny", z. 4, s. 303-315.

Wiśniewski J.W. (2009), *Mikroekonometria*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń.

Wiśniewski J.W. (2012), *Dilemmas of Economic Measurements in Weak Scales*, "Folia Oeconomica Stetinensia", No. 10(18), s. 50-59.

Wiśniewski J.W. (2013a), *Forecasting Staffing Decisions*, *EKONOMETRIA. ECONOMETRICS* 1(39), Publishing House of Wrocław, University of Economics Wrocław, s. 22-29.

Wiśniewski J.W. (2013b), *Correlation and Regression of Economic Qualitative Features*, LAP LAMBERT Academic Publishing, Saarbrücken.

Wiśniewski J.W. (2016), *Microeconometrics in Business Management*, John Wiley&Sons, Chichester.

LINEAR FUNCTION OF PROBABILITY

Summary: An important analytical tool in econometrics, serving, inter alia, to investigate the association of dummy variables, is a linear function of probability. The function is also known as Goldberger model. Its specificity is the dummy dependent variable which causes, that theoretical values of the empirical model are empirical estimates of probability of a variant, signed as 1. The explanatory variables in the model can be both continuous or discrete. The Goldberger model is an important instrument for the measurement of the casual conditions, mainly qualitative variables, but also the quantitative variables. However, it requires a specific approach, primarily in estimating of its parameters.

Keywords: dummy variable, association, probability, econometric model.