

Beata Bieszk-Stolorz

Uniwersytet Szczeciński

WYKORZYSTANIE MODELU LOGITOWEGO DO ANALIZY BEZROBOCIA WŚRÓD OSÓB NIEPEŁNOSPRAWNYCH W POLSCE W 2010 ROKU

Streszczenie

Celem artykułu było zbadanie, w jaki sposób wybrane zmienne: płeć, miejsce zamieszkania, wiek, wykształcenie i stopień niepełnosprawności, wpłynęły na bezrobocie wśród osób niepełnosprawnych w Polsce w 2010 roku. W przeprowadzonej analizie zastosowano dwumianowy i wielomianowy model regresji logistycznej. Oszacowane parametry tych modeli umożliwiły wyznaczenie ilorazów ryzyka bezrobocia w stosunku do posiadania pracy przez osoby niepełnosprawne, a także pozwoliły obliczyć prawdopodobieństwo przynależności do kategorii bezrobotnych w zależności od płci, miejsca zamieszkania osoby niepełnosprawnej oraz prawdopodobieństwo długości okresu poszukiwania pracy w zależności od płci.

Słowa kluczowe: niepełnosprawność, bezrobocie, modele regresji logistycznej.

Wprowadzenie

Liczba osób niepełnosprawnych w Polsce w 2002 roku wynosiła prawie 5,5 mln, w tym około 4,5 mln miało prawne potwierdzenie faktu niepełnosprawności, a 4,3 mln stanowiły osoby w wieku 15 lat i więcej [4]. Są to dane z Narodowego Spisu Powszechnego z 2002 roku. Liczba osób niepełnosprawnych w wieku produkcyjnym z orzeczoną niepełnosprawnością systematycznie

spadała i w 2010 roku wyniosła 3393 tys. osób, co stanowiło 10,7 % ludności w wieku 15 lat i więcej [1]. W ostatnich latach obserwuje się wzrost liczby zatrudnionych niepełnosprawnych. W roku 2010 współczynnik aktywności zawodowej tych osób wyniósł 17,4%, stopa bezrobocia (13,6%) była wyższa niż stopa bezrobocia liczona dla wszystkich osób w analogicznym okresie (9,3%). Celem artykułu jest zbadanie, w jaki sposób wybrane zmienne: płeć, miejsce zamieszkania, wiek, wykształcenie i stopień niepełnosprawności, wpłynęły na bezrobocie wśród osób niepełnosprawnych w Polsce w 2010 roku. W przeprowadzonej analizie zastosowano dwumianowy i wielomianowy model regresji logistycznej. Oszacowane parametry tych modeli umożliwiły wyznaczenie ilorazów ryzyka bezrobocia w stosunku do posiadania pracy przez osoby niepełnosprawne, a także obliczyć prawdopodobieństwo przynależności do kategorii bezrobotnych w zależności od płci, miejsca zamieszkania osoby niepełnosprawnej oraz prawdopodobieństwo długości okresu poszukiwania pracy w zależności od płci.

1. Dane statystyczne wykorzystane w analizie

Dane liczbowe zaczerpnięto z publikacji GUS *Aktywność ekonomiczna ludności Polski, IV kwartał 2010*. Za ludność aktywną ekonomicznie przyjęto ludność aktywną lub bierną zawodowo. Ludność aktywna zawodowo obejmuje wszystkie osoby pracujące (powyżej 15 lat i więcej) lub bezrobotne. Populacja osób bezrobotnych obejmuje osoby w wieku od 15 do 74 lat, spełniające jednocześnie trzy warunki: w okresie badanego tygodnia nie były osobami pracującymi, aktywnie poszukiwały pracy, były gotowe (zdolne) podjąć pracę w ciągu dwóch tygodni następujących po tygodniu badanym. Do bezrobotnych zaliczono również osoby, które nie poszukiwały pracy, ponieważ miały załatwioną pracę oraz były gotowe ją podjąć. Do osób niepełnosprawnych zaliczono osoby w wieku 16 lat i więcej, które mają orzeczenie o stopniu niepełnosprawności lub niezdolności do pracy. Dane dotyczyły płci, wieku, wykształcenia, miejsca zamieszkania i stopnia niepełnosprawności.

Niepełnosprawność badanych osób podzielono na trzy grupy. Pierwszy stopień niepełnosprawności (znaczny) dotyczy osób mających przyznaną pierwszą grupę inwalidzką lub orzeczenie o całkowitej niezdolności do pracy i samodzielnej egzystencji. Drugi stopień niepełnosprawności (umiarkowany) dotyczy

osób mających przyznaną drugą grupę inwalidzką lub orzeczenie o całkowitej niezdolności do pracy. Do trzeciego stopnia niepełnosprawności (lekkiego) zaliczono osoby mające przyznaną trzecią grupę inwalidzką lub orzeczenie o częściowej niezdolności do pracy lub orzeczenie o niezdolności do pracy w gospodarstwie rolnym. Analizie poddano ogółem 590 tys. osób niepełnosprawnych aktywnych zawodowo, a ich strukturę podano w tabeli 1.

Tabela 1. Struktura osób niepełnosprawnych aktywnych zawodowo w wieku 16 lat i więcej w Polsce w 2010 roku (tys.)

Cechy	Grupy (oznaczenie)	Aktywni zawodowo	
		razem	w tym bezrobotni
Płeć	kobiety (K)	242	30
	mężczyźni (M)	348	47
Miejsce zamieszkania	miasto (T)	385	57
	wieś (W)	205	23
Wykształcenie	co najwyżej gimnazjalne (W1)	90	15
	zasadnicze zawodowe (W2)	216	35
	średnie ogólnokształcące (W3)	40	5
	policealne i średnie zawodowe (W4)	165	21
	wyższe (W5)	80	5
Wiek	15–24 (S1)	20	6
	25–34 (S2)	78	17
	35–44 (S3)	85	11
	45–54 (S4)	190	28
	55–64 (S5)	181	18
	powyżej 65 (S6)	37	–
Stopień niepełnosprawności	I stopień (I)	53	6
	II stopień (II)	232	34
	III stopień (III)	306	39

Źródło: opracowanie własne na podstawie danych z GUS.

Ponieważ wśród badanych osób niepełnosprawnych nie było osób bezrobotnych powyżej 65. roku życia, badaną zbiorowość podzielono według płci i miejsca zamieszkania na dwie grupy, według wykształcenia i wieku na pięć grup, a według stopnia niepełnosprawności na trzy grupy. Spośród wszystkich aktywnych ekonomicznie osób niepełnosprawnych 82,6% stanowiły osoby bierne zawodowo, a wśród aktywnych zawodowo osób niepełnosprawnych 13,6% – to osoby bezrobotne.

2. Model regresji logistycznej

Do analizy danych wykorzystano model regresji logistycznej [2]. Funkcja logistyczna ma postać:

$$P(Z) = \frac{\exp Z}{1 + \exp Z} \quad (1)$$

Przyjmuje ona wartości od 0 do 1 i wyraża prawdopodobieństwo wystąpienia określonego zdarzenia Z . W przypadku zmiennej objaśnianej dychotomicznej model ten można zapisać następująco [3]:

$$P(Y = 1 | x_1, x_2, \dots, x_k) = \frac{\exp\left(\alpha_0 + \sum_{i=1}^k \alpha_i x_i\right)}{1 + \exp\left(\alpha_0 + \sum_{i=1}^k \alpha_i x_i\right)} \quad (2)$$

gdzie:

- Y – dychotomiczna zmienna zależna,
- x_1, x_2, \dots, x_k – zmienne niezależne,
- α_i – współczynniki regresji.

Zapis $p = P(Y = 1)$ określa prawdopodobieństwo sukcesu (lub porażki), wyrażenie $\frac{p}{1-p}$ oznacza szansę (lub ryzyko) wystąpienia określonego zdarzenia.

Wyrażenie $\ln\left(\frac{p}{1-p}\right)$ zapisuje się jako $\text{logit}(p)$ i stosuje w zapisie modelu logitowego:

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \alpha_0 + \sum_{i=1}^k \alpha_i x_i \quad (3)$$

Do analizy oszacowanych parametrów wykorzystuje się ich przekształconą postać $\exp(\alpha_i)$ nazywaną ilorazem szans (*odds ratio* – OR). W przypadku zmiennej objaśnianej, która nie jest dychotomiczna i ma wartości ze zbioru $\{0, 1, 2, \dots, s\}$ można zastosować wielomianowy model logistyczny [4], który ma następującą postać:

$$P(Y = 0 | x_1, x_2, \dots, x_m) = \frac{1}{1 + \sum_{j=1}^s \exp\left(\alpha_{j0} + \sum_{i=1}^m \alpha_{ji}x_i\right)} \quad (4)$$

$$P(Y = j | x_1, x_2, \dots, x_m) = \frac{\exp\left(\alpha_{j0} + \sum_{i=1}^m \alpha_{ji}x_i\right)}{1 + \sum_{j=1}^s \exp\left(\alpha_{j0} + \sum_{i=1}^m \alpha_{ji}x_i\right)} \quad \text{dla } j = 1, 2, \dots, s \quad (5)$$

gdzie:

- j – numer wariantu zmiennej objaśnianej, $j = 0, 1, \dots, s$,
- i – numer zmiennej objaśniającej, $i = 1, 2, \dots, m$.

3. Analiza bezrobocia wśród osób niepełnosprawnych

W pierwszym etapie analizy zbadano wpływ płci, miejsca zamieszkania, wykształcenia, wieku oraz stopnia niepełnosprawności na bezrobocie wśród osób niepełnosprawnych. Dychotomiczną zmienną zależną (aktywność zawodowa) zakodowano w następujący sposób: bezrobotni – 1, pracujący – 0. Do kodowania zmiennych objaśniających zastosowano kodowanie 0-1, które umożliwiło porównanie wyodrębnionych grup poszczególnych cech (zakodowanych jako 1) z wybraną grupą (zakodowaną jako 0) [3]. Dla płci są to mężczyźni, dla miejsca zamieszkania – wieś, dla grup wykształcenia – wykształcenie co najwyżej gimnazjalne, dla grup wieku – przedział wieku 15–24, dla grup stopnia niepełnosprawności – pierwszy stopień niepełnosprawności. W tabeli 2 przedstawiono wyniki szacowania parametrów dwumianowych modeli logitowych. Ze względu na badane zjawisko (bezrobocie) autorka zamiast określenia szansa będzie używała pojęcia ryzyko.

Z analizy ilorazu szans (ryzyka), przedstawionych na rysunku 1, wynika, że największe ryzyko bezrobocia dotyczyło mężczyzn, osób niepełnosprawnych mieszkających w mieście, osób z wykształceniem co najwyżej gimnazjalnym, z drugim stopniem niepełnosprawności, w wieku od 15 do 24 lat.

Tabela 2. Wyniki estymacji parametrów modeli logitowych aktywności ekonomicznej względem cech osób niepełnosprawnych

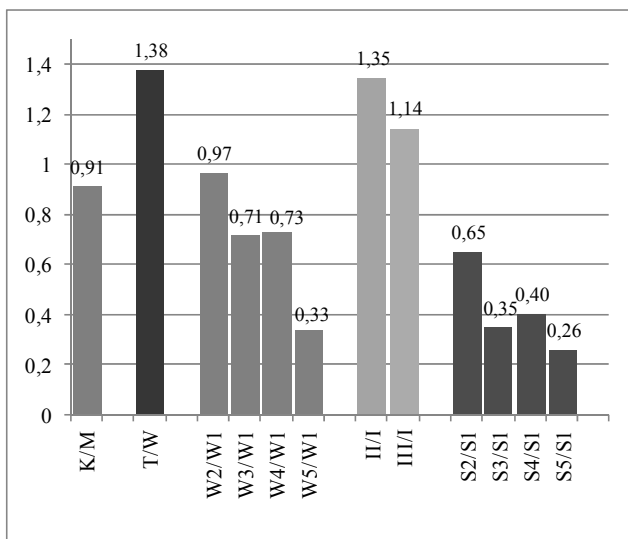
Zmienne	Ocena	Błąd standardowy	p	Iloraz ryzyka
Płeć ($\chi^2 = 149,59, p = 0,0000$)				
Stała	-1,8570	0,0050	0,0000	
K/M	-0,0931	0,0076	0,0000	0,9111
Miejsce zamieszkania ($\chi^2 = 1506, p = 0,0000$)				
Stała	-2,06851	0,006998	0,0000	
T/W	0,31855	0,008341	0,0000	1,3751
Wykształcenie ($\chi^2 = 6482, p = 0,0000$)				
Stała	-1,6094	0,0089	0,0000	
W2/W1	-0,0337	0,0107	0,0000	0,9669
W3/W1	-0,3365	0,0176	0,0000	0,7143
W4/W1	-0,3159	0,0116	0,0000	0,7292
W5/W1	-1,0986	0,0171	0,0000	0,3333
Rodzaj niepełnosprawności ($\chi^2 = 628, p = 0,0000$)				
Stała	-2,0584	0,0137	0,0000	
II/I	0,2965	0,0149	0,0000	1,3451
III/I	0,1347	0,0147	0,0000	1,1442
Wiek ($\chi^2 = 9625, p = 0,0000$)				
Wyraz wolny	-0,8473	0,0154	0,0000	
S2/S1	-0,4304	0,0177	0,0000	0,6503
S3/S1	-1,0589	0,0185	0,0000	0,3468
S4/S1	-0,9081	0,0167	0,0000	0,4033
S5/S1	-1,3561	0,0173	0,0000	0,2577

Źródło: opracowanie własne.

W drugim etapie analizy, korzystając z wielomianowych modeli logitowych, zbadano prawdopodobieństwo przynależności do kategorii bezrobotnych w zależności od płci, miejsca zamieszkania osoby niepełnosprawnej oraz obliczono prawdopodobieństwo długości okresu poszukiwania pracy w zależności od płci.

W pierwszej grupie modeli bezrobotnych niepełnosprawnych wyodrębniono trzy kategorie (zmienna objaśniana): osoby, które straciły pracę, osoby powracające do pracy po przerwie, osoby podejmujące pracę po raz pierwszy. Wyniki szacowania parametrów przedstawiono w tabeli 3. Najbardziej prawdopodobną przyczyną pozostawania bezrobotnym wśród osób niepełnosprawnych we wszystkich badanych grupach było podjęcie pracy po raz pierwszy (tabela 4), przy czym dla kobiet prawdopodobieństwo to było największe i równe 0,59. Drugą ważną przyczyną we wszystkich grupach był powrót do pracy po przerwie.

Rysunek 1. Ryzyko względne bezrobocia wśród osób niepełnosprawnych według płci, miejsca zamieszkania, wykształcenia, stopnia niepełnosprawności i wieku w Polsce w 2010 roku



Źródło: opracowanie własne.

Tabela 3. Wyniki estymacji wielomianowych modeli logitowych przynależności do kategorii bezrobotnych według płci i miejsca zamieszkania

Zmienne	Ocena	Błąd standardowy	<i>p</i>
Płeć			
Wyraz wolny 1	1,0296	0,0165	0,0000
K/M	-0,1542	0,0236	0,0000
Wyraz wolny 2	1,6864	0,0154	0,0000
K/M	-0,6568	0,0225	0,0000
Miejsce zamieszkania			
Wyraz wolny 1	0,5596	0,0198	0,0000
T/W	0,4389	0,0243	0,0000
Wyraz wolny 2	1,0116	0,0185	0,0000
T/W	0,4437	0,0227	0,0000

Źródło: opracowanie własne.

Tabela 4. Ocena prawdopodobieństwa przynależności do kategorii bezrobotnych osób niepełnosprawnych w Polsce w 2010 roku

Cecha	Utrata pracy	Powrót do pracy po przerwie	Podjęcie pracy po raz pierwszy
Kobieta	0,11	0,30	0,59
Mężczyzna	0,16	0,39	0,45
Miasto	0,18	0,32	0,50
Wieś	0,13	0,34	0,54

* pogrubioną czcionką zaznaczono największe wartości prawdopodobieństwa.

Źródło: opracowanie własne.

W kolejnym, wielomianowym modelu logitowym zmienną objaśnianą był okres poszukiwania pracy. W tym przypadku wyróżniono cztery kategorie bezrobotnych: do trzech miesięcy włącznie, od czterech do sześciu miesięcy, od siedmiu do dwunastu miesięcy oraz powyżej dwunastu miesięcy. Wyniki estymacji parametrów przedstawiono w tabeli 5.

Tabela 5. Wyniki estymacji parametrów wielomianowego modelu logitowego okresu poszukiwania pracy przez bezrobotnych niepełnosprawnych według płci

Zmienne	Ocena	Błąd standardowy	<i>p</i>
Płeć			
Wyraz wolny 1	-0,2683	0,0117	0,0000
K/M	-0,2914	0,0182	0,0000
Wyraz wolny 2	-0,8873	0,0142	0,0000
K/M	0,3277	0,0200	0,0000
Wyraz wolny 3	-0,6360	0,0130	0,0000
K/M	-0,3936	0,0210	0,0000

Źródło: opracowanie własne.

Tabela 6. Ocena prawdopodobieństwa długości okresu poszukiwania pracy przez osoby niepełnosprawne w Polsce w 2010 roku

Cecha	0–3 miesiące	4–6 miesięcy	7–12 miesięcy	Powyżej 12 miesięcy
Kobieta	0,40	0,23	0,23	0,14
Mężczyzna	0,37	0,28	0,15	0,20

* pogrubioną czcionką zaznaczono największe wartości prawdopodobieństwa

Źródło: opracowanie własne.

Najbardziej prawdopodobny okres poszukiwania pracy przez bezrobotnych niepełnosprawnych mężczyzn (0,37) i kobiety (0,40) wynosi do trzech miesięcy łącznie (tabela 6), a w drugiej kolejności od czterech do sześciu miesięcy.

Podsumowanie

W roku 2010 płeć, miejsce zamieszkania, wiek, wykształcenie i stopień niepełnosprawności były determinantami bezrobocia osób niepełnosprawnych w Polsce. W tym okresie większą szansę na pracę miały kobiety, osoby niepełnosprawne mieszkające na wsi, z wykształceniem wyższym, z pierwszym stopniem niepełnosprawności, w wieku od 60 do 65 roku życia. Najbardziej prawdopodobną przyczyną pozostawania bez pracy, bez względu na płeć i miejsce zamieszkania, była próba podjęcia pracy po raz pierwszy, czyli moment wkraczania na rynek pracy. Pozytywnym zjawiskiem jest fakt, że najbardziej prawdopodobnym okresem poszukiwania pracy przez osoby niepełnosprawne (zarówno przez kobiety, jak i mężczyzn) był najkrótszy okres, czyli do trzech miesięcy łącznie, co może świadczyć o skuteczności programów zachęcających pracodawców do zatrudniania takich osób.

Literatura

1. *Aktywność ekonomiczna ludności Polski, IV kwartał 2010*, Informacje i Opracowania Statystyczne, GUS, Warszawa 2011.
2. Cramer J.S., *The Origins of Logistic Regression*, Tinbergen Institute Discussion Paper, Faculty of Economics and Econometrics, University of Amsterdam, Amsterdam 2002.
3. Hosmer D.W., Lemeshow S., *Applied Logistic Regression*, John Wiley & Sons Inc., New York 2000.
4. Kleinbaum D.G., Klein M., *Logistic Regression. A Self-Learning Text*. second edit., Springer-Verlag, New York–Berlin–Heidelberg 2002.
5. *Niepełnosprawność w liczbach*, <http://www.niepelnosprawni.gov.pl/niepelnospraw-nosc-w-liczbach/>, dostęp 15.07.2011.

THE APPLICATION OF THE LOGIT MODEL IN ANALYSING UNEMPLOYMENT AMONG THE DISABLED PERSONS IN POLAND IN 2010

Summary

The aim of the paper was to examine how the chosen variables, such as gender, the place of residence, age, education and the level of disability influenced the rate of unemployment among the disabled persons in Poland in 2010. The authors applied the binomial and the multinomial logistic regression model. The estimated parameters of these models helped them to determine the unemployment odds rates in relation to the disabled persons' employment as well as to calculate the likelihood of their affiliation to the category of unemployed people depending on a disabled person's gender and place of residence. Additionally, the authors determined the likelihood of the job seeking period depending on gender.

Keywords: disabled persons, unemployment, logistic regression models.

Translated by Anita Zdrojewska