

**Barbara Batóg, Magdalena Mojsiewicz**

Uniwersytet Szczeciński

**Katarzyna Wawrzyniak**

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

---

## **KLASYFIKACJA GOSPODARSTW DOMOWYCH POD WZGLĘDEM POPYTU POTENCJALNEGO I ZREALIZOWANEGO NA RYNKU UBEZPIECZEŃ W POLSCE**

---

**Streszczenie:** W artykule omówiono problematykę popytu potencjalnego i zrealizowanego na ubezpieczenia majątkowe, na życie, na dożycie oraz zdrowotne. Dokonano klasyfikacji gospodarstw domowych przy wykorzystaniu liniowych modeli regresji ze zmiennymi parametrami. Zmiennymi objaśnianymi dla popytu potencjalnego były deklarowane składki miesięczne i ich transformacje, a dla popytu zrealizowanego liczba polis wykupionych na ubezpieczenia danego typu. Podstawą do wyodrębnienia grup były kategorie zmiennych jakościowych.

### **1. Wstęp**

Celem badania była klasyfikacja gospodarstw domowych na rynku ubezpieczeń w Polsce pod względem potencjalnego i zrealizowanego popytu na ubezpieczenia majątkowe, na życie, na dożycie oraz pokrywające koszty leczenia. Popyt potencjalny reprezentowany jest przez deklarowaną miesięczną składkę na wymienione rodzaje ubezpieczeń oraz jej transformacje, a popyt zrealizowany przez liczbę posiadanych polis danego typu. Do klasyfikacji wykorzystano modele regresji ze zmiennymi parametrami. Klasyfikacje przeprowadzono dla reprezentatywnej próby 500 gospodarstw domowych<sup>1</sup>.

W artykule postawiono hipotezę o zbieżności zbioru zmiennych objaśniających w przypadku modelowania popytu zrealizowanego i popytu potencjalnego na rynku ubezpieczeń w Polsce. Dodatkowo weryfikowano hipotezę o większej przydatności oceny zobiektywizowanej niż oceny subiektywnej popytu potencjalnego

---

<sup>1</sup> Próba liczyła 500 gospodarstw domowych, a badania terenowe zostały przeprowadzone w lipcu i sierpniu 2005 r. przez TNS OBOP w ramach grantu KBN „Metody statystyczne w segmentacji rynku ubezpieczeń w Polsce” nr 1H02B02827.

w ubezpieczeniach. Ocenę zobiektywizowaną otrzymano przez transformację zmiennej *deklarowana składka*, odnoszącej się do hipotetycznych warunków zakupu ubezpieczenia. Transformacja ta okazała się możliwa po pierwsze w przypadku ubezpieczeń, w których beneficjentem są spadkobiercy tożsami z członkami gospodarstwa domowego według stanu z okresu badawczego. Powiązано wówczas deklarowaną składkę na ubezpieczenie klasyczne na życie z prawdopodobieństwem zgonu w ciągu roku osoby ubezpieczającej, uwzględniając jej wiek, płeć i miejsce zamieszkania, jak również z obecnymi dochodami gospodarstwa domowego. Druga transformacja dotyczyła *deklarowanej składki* w ubezpieczeniach zapewniających dochody po przekroczeniu pewnego progu wieku, a więc zabezpieczających wprost osobę ubezpieczającą. Ta transformacja wiązała deklarowaną składkę z prawdopodobieństwem przeżycia roku przez osobę badaną, uwzględniając jej płeć, wiek i miejsce zamieszkania z obecnymi przychodami gospodarstwa domowego.

## 2. Metodologia wykorzystana w badaniu

W celu wyodrębnienia kategorii zmiennych różnicujących w sposób istotny gospodarstwa domowe pod względem potencjalnego i zrealizowanego popytu wykorzystano liniowe modele regresji ze zmiennymi parametrami. W tego typu modelach każda ze zmiennych objaśniających jakościowych jest reprezentowana przez zmienne zero-jedynkowe zdefiniowane jako [Johnston 1991; Maddala 2006]:

$$D_{lk} = \begin{cases} 1 & \text{dla kategorii } k \\ 0 & \text{dla innych kategorii} \end{cases}$$

gdzie:  $l$  – numer zmiennej jakościowej,  
 $k = 2, \dots, K$ ,  
 $K$  – liczba kategorii zmiennej jakościowej  $l$ .

Po oszacowaniu modelu parametry stojące przy zmiennych  $D_{lk}$  reprezentują odchylenia od poziomu kategorii bazowej dla  $k = 1$ . Jeżeli parametr przy zmiennej  $D_{lk}$  jest statystycznie istotny, to przypadki przyporządkowane kategorii  $k$  różnią się co do wartości przyjmowanych przez zmienną objaśnianą dla przypadków z kategorii bazowej ( $k = 1$ ).

## 3. Wyniki estymacji liniowych modeli regresji popytu potencjalnego

Dla popytu potencjalnego oszacowano 5 liniowych modeli regresji ze zmiennymi parametrami. Zmiennymi objaśnianymi były odpowiednio:

- dla ubezpieczeń na życie: deklarowana miesięczna składka na tego typu ubezpieczenia (*składka życie*) oraz jej transformacja – zmienna  $y_1$ ,

- dla ubezpieczeń na dożycie (spełniających funkcję zabezpieczeń emerytalnych): deklarowana miesięczna składka na tego typu ubezpieczenia (*składka dożycie*) oraz jej transformacja – zmienna  $y_2$ ,
- dla ubezpieczeń zdrowotnych (ubezpieczeń od kosztów leczenia): deklarowana miesięczna składka na tego typu ubezpieczenia (*składka zdrowie*).

Zmienna  $y_1$  została wyznaczona według wzoru:

$$y_1 = \frac{dsm_1}{{}_1q_x} : \omega_1,$$

gdzie:  $dsm_1$  – zadeklarowana składka miesięczna na ubezpieczenie na życie,  
 ${}_1q_x$  – prawdopodobieństwo zgonu w ciągu roku osoby w wieku  $x$  lat;  
 $\omega_1$  – miesięczny dochód rozporządzalny gospodarstwa domowego.

Zmienna  $y_2$  została wyznaczona według wzoru:

$$y_2 = \frac{dsm_2}{{}_k p_x} : \omega_2,$$

gdzie:  $dsm_2$  – zadeklarowana składka miesięczna na ubezpieczenie na dożycie,  
 $x+k$  – wiek emerytalny,  
 ${}_k p_x$  – prawdopodobieństwo dożycia osoby w wieku  $x$  lat do wieku emerytalnego,  
 $\omega_2$  – miesięczny dochód rozporządzalny gospodarstwa domowego przypadający na jedną osobę dorosłą.

Wprowadzenie zmiennych  $y_1$  oraz  $y_2$  wynikało z faktu, że w przypadku ubezpieczeń życiowych składka powinna być związana z relacją subiektywnej wyceny świadczenia do obecnych przeciętnych dochodów rozporządzalnych osób w gospodarstwie domowym, przy założeniu, że ubezpieczenia na życie i na dożycie mają chronić bieżące dochody rozporządzalne<sup>2</sup>. Zgodnie z wcześniejszym opisem wprowadzone transformacje były próbą zobiektywizowania wyceny zaproponowanej przez respondentów.

W modelowaniu nie uwzględniono zmiennej *deklarowana składka* w ubezpieczeniach majątku, choć badanie zostało przygotowane tak, aby taką procedurę przeprowadzić. Respondenci wskazywali wysokość składki, którą skłonni byliby zapłacić za ubezpieczenie samochodu oraz ubezpieczenie mieszkania przy założonych z góry prawdopodobieństwach realizacji zdarzenia polegającego na zajściu szkody całkowitej. W pracach [Mojsiewicz 2008; Batóg, Mojsiewicz, Wawrzyniak 2007] wykazano, że otrzymane wyniki można jedynie transponować na zmienne wyrażone na skali porządkowej.

W zbiorze zmiennych objaśniających dla każdego typu ubezpieczeń znalazły się następujące zmienne charakteryzujące gospodarstwo domowe: wiek, wykształ-

<sup>2</sup> To uproszczenie jest przyjęte celowo, by zaprezentować jedynie ideę porządkowania ryzyka.

cenie głowy gospodarstwa domowego, miejsce zamieszkania, liczba osób w gospodarstwie domowym, poziom posiadanych oszczędności. Dla potrzeb estymacji modeli regresji popytu potencjalnego ze zmiennymi parametrami dla jakościowych zmiennych objaśniających w następujący sposób zdefiniowano kategorie bazowe:

- wykształcenie głowy gospodarstwa domowego: kategoria bazowa „co najwyżej gimnazjalne”, zmienne zero-jedynkowe dla kolejnych kategorii oznaczono jako: W2 – zasadnicze zawodowe, W3 – średnie, W4 – wyższe,
- miejsce zamieszkania: kategoria bazowa „wieś”, zmienne zero-jedynkowe dla kolejnych kategorii oznaczono jako: A2 – miasto o liczbie ludności 50 tys. i mniej, A3 – miasto o liczbie ludności 51-200 tys., A4 – miasto o liczbie ludności 201-500 tys., A5 – miasto o liczbie ludności powyżej 500 tys.,
- poziom posiadanych oszczędności: kategoria bazowa „brak oszczędności”, zmienne zero-jedynkowe dla kolejnych kategorii oznaczono jako: osz2 – w wysokości nie wyższej niż miesięczne dochody rodziny, osz3 – w wysokości pomiędzy miesięcznymi a rocznymi dochodami rodziny, osz4 – w wysokości co najmniej rocznych dochodów rodziny.

Wyniki estymacji liniowych modeli regresji ze zmiennymi parametrami dla popytu potencjalnego przedstawiono w tab. 1. Parametry istotne statystycznie zostały pogrubione.

**Tabela 1.** Wyniki estymacji liniowych modeli regresji ze zmiennymi parametrami dla popytu potencjalnego

Zmienne objaśniające	Zmienna objaśniana				
	składka życie	$y_1$	składka dożycie	$y_2$	składka zdrowie
Wyraz wolny	<b>-64,808</b>	<b>11,828</b>	<b>-17,628</b>	-33,636	-6,305
Wiek	<b>0,579</b>	<b>-0,192</b>	-0,004	<b>12,305</b>	-0,078
W2	-0,377	<b>-1,749</b>	-1,314	<b>349,939</b>	-6,535
W3	<b>39,919</b>	-0,778	<b>7,634</b>	<b>890,033</b>	<b>10,731</b>
W4	<b>-47,091</b>	<b>-5,184</b>	<b>-29,640</b>	<b>1 240,866</b>	<b>75,084</b>
A2	<b>-12,944</b>	0,125	<b>-8,617</b>	-14,699	<b>-19,081</b>
A3	-11,225	0,091	0,412	<b>190,218</b>	<b>-10,535</b>
A4	<b>-17,438</b>	-1,161	0,473	<b>-535,845</b>	<b>-15,533</b>
A5	<b>58,462</b>	<b>5,741</b>	<b>45,294</b>	<b>-611,419</b>	<b>10,816</b>
Liczba osób w gosp. domowym	<b>15,576</b>	-0,001	<b>7,096</b>	<b>-74,449</b>	<b>8,483</b>
osz2	88,563	2,417	36,489	448,584	-9,059
osz3	<b>40,428</b>	-0,739	2,709	<b>637,972</b>	<b>21,701</b>
osz4	<b>-33,272</b>	2,079	4,709	83,279	3,098
R <sup>2</sup>	0,373	0,323	0,462	0,552	0,458
F(12,487)	24,183	19,379	34,913	49,990	34,231
P	0	0	0	0	0

Źródło: obliczenia własne.

Okazało się, że dla wszystkich zmiennych objaśnianych istotne były parametry stojące przy zmiennych W4 oraz A5. Oznacza to, że gospodarstwa domowe, gdzie głowa rodziny miała wykształcenie wyższe, oraz zamieszkujące miasto powyżej 500 tys. mieszkańców charakteryzowały się innym poziomem zmiennych objaśnianych niż gospodarstwa domowe przynależące do pozostałych kategorii zmiennych *wykształcenie* oraz *miejsce zamieszkania*. Z kolei parametry stojące przy zmiennej *osz2* były w przypadku każdej zmiennej objaśnianej nieistotne statystycznie. Oznacza to, że gospodarstwa domowe, u których wystąpił brak oszczędności, oraz gospodarstwa domowe, które posiadały oszczędności w wysokości nie wyższej niż miesięczne dochody rodziny, charakteryzowały się praktycznie tym samym poziomem zmiennych objaśnianych.

W modelach dla ubezpieczeń na życie i na dożycie wprowadzenie zmiennych sztucznych  $y_1$  i  $y_2$  (składka po zobiektywizowaniu) poprawiło wyniki estymacji jedynie w przypadku ubezpieczeń mających charakter zabezpieczeń emerytalnych (transformacja  $y_2$ ):  $R^2$  wyższe o ok. 9% i większy zbiór zmiennych objaśniających: 9 zmiennych w stosunku do wcześniejszych 5 zmiennych. Druga transformacja ( $y_1$ ) pogorszyła dopasowanie modelu, jak również doprowadziła do redukcji zbioru zmiennych objaśniających statystycznie istotnych.

#### 4. Wyniki estymacji liniowych modeli regresji popytu zrealizowanego

Dla popytu zrealizowanego oszacowano 3 liniowe modele regresji ze zmiennymi parametrami. Zmiennymi objaśnianymi były odpowiednio:

- dla ubezpieczeń majątkowych: łączna liczba polis majątkowych posiadanych przez gospodarstwo domowe,
- dla ubezpieczeń na życie: łączna liczba polis na życie posiadanych przez gospodarstwo domowe,
- dla ubezpieczeń na dożycie (zabezpieczenie emerytalne): łączna liczba polis na dożycie posiadanych przez gospodarstwo domowe.

W okresie prowadzenia badania ubezpieczenia zdrowotne miały jeszcze formę abonamentów medycznych, a pytanie o *deklarowaną składkę* było postawione w odniesieniu do sytuacji hipotetycznej (wyłącznie popytu potencjalnego). Zbiór zmiennych objaśniających w przypadku popytu zrealizowanego był taki sam jak w przypadku popytu potencjalnego.

Wyniki estymacji liniowych modeli regresji ze zmiennymi parametrami dla popytu potencjalnego przedstawiono w tab. 2. Parametry istotne statystycznie zostały pogrubione.

Okazało się, że dla wszystkich rodzajów ubezpieczeń istotne były parametry stojące przy zmiennych W3, A2, A3, A4, liczba osób w gospodarstwie domowym oraz *osz3*. Istotność zmiennej W3 oznacza, że gospodarstwa domowe, w których głowa gospodarstwa domowego ma wykształcenie średnie, charakteryzują się posiadaniem innej liczby polis niż pozostałe gospodarstwa domowe. Istotność zmien-

**Tabela 2.** Wyniki estymacji liniowych modeli regresji ze zmiennymi parametrami dla popytu zrealizowanego

Zmienne objaśniające	Rodzaj ubezpieczenia		
	majątkowe	na życie	na dożycie
Wyraz wolny	0,450	0,555	<b>-0,759</b>
Wiek	0,005	-0,002	<b>0,014</b>
W2	<b>0,653</b>	-0,084	<b>-0,363</b>
W3	<b>0,567</b>	<b>0,831</b>	<b>0,636</b>
W4	<b>1,803</b>	0,026	<b>0,567</b>
A2	<b>-0,578</b>	<b>-0,806</b>	<b>-0,565</b>
A3	<b>-1,110</b>	<b>-0,757</b>	<b>-0,780</b>
A4	<b>-1,068</b>	<b>-1,337</b>	<b>-0,317</b>
A5	<b>-0,923</b>	0,128	<b>-0,514</b>
Liczba osób w gosp. domowym	<b>0,133</b>	<b>0,418</b>	<b>0,221</b>
osz2	1,054	0,959	0,232
osz3	<b>2,707</b>	<b>3,068</b>	<b>2,970</b>
osz4	-0,041	0,736	<b>1,303</b>
R <sup>2</sup>	0,498	0,309	0,496
F(12,487)	40,144	18,144	39,881
P	0	0	0

Źródło: obliczenia własne.

nych A2, A3, A4 oznacza, że gospodarstwa domowe zamieszkujące miasta, liczące od 50 do 500 tys. mieszkańców, mają inną liczbę polis niż gospodarstwa domowe na wsi. Liczba osób w gospodarstwie domowym istotnie wpływa na liczbę polis każdego z rozważanych rodzajów ubezpieczeń posiadanych przez gospodarstwo domowe. Istotność zmiennej osz3 oznacza, że gospodarstwa domowe, które mają oszczędności w wysokości pomiędzy miesięcznymi a rocznymi dochodami rodziny, mają inną liczbę polis niż pozostałe gospodarstwa.

Podobnie jak w przypadku popytu potencjalnego parametry stojące przy zmiennej osz2 były w przypadku każdej zmiennej objaśnianej nieistotne statystycznie. Oznacza to, że gospodarstwa domowe, w których wystąpił brak oszczędności, oraz gospodarstwa domowe, które miały oszczędności w wysokości nie wyższej niż miesięczne dochody rodziny, charakteryzowały się praktycznie tą samą liczbą zarówno polis majątkowych, jak i polis na życie oraz dożycie.

## 5. Wnioski

W przypadku ubezpieczeń na życie (związanych z zabezpieczeniem wskazanych członków gospodarstwa domowego), gdy porówna się modele regresji dla zmiennej objaśnianej *składka na życie w ocenie subiektywnej* oraz zmiennej objaś-

nianej liczba polis na życie posiadanych przez gospodarstwo domowe, można zauważyć, że:

- wiek był zmienną różnicującą gospodarstwa domowe tylko w modelu popytu potencjalnego,
- zmienna reprezentująca kategorię W3 (wykształcenie średnie) różnicowała gospodarstwa domowe w obydwu modelach przy dodatnich oszacowaniach parametrów (z różną siłą),
- zmienne reprezentujące kategorię A2 (miasto o liczbie ludności 50 tys. i mniej) i A4 (miasto o liczbie ludności 201-500 tys.) w obydwu modelach charakteryzowały się znakiem ujemnym (z różną siłą),
- liczba osób w gospodarstwie domowym oraz kategoria oszcz3 (w wysokości pomiędzy miesięcznymi a rocznymi dochodami rodziny) charakteryzowały się dodatnim oszacowaniem parametrów w obydwu modelach.

W przypadku ubezpieczeń na dożycie (związanych z zabezpieczeniem emerytalnym), gdy porówna się modele regresji dla zmiennej objaśnianej *zobiektywizowany popyt potencjalny* oraz zmiennej objaśnianej *popyt zrealizowany mierzony liczbą posiadanych polis ubezpieczeń na dożycie*, okazuje się, że:

- wiek w obydwu modelach charakteryzował się dodatnim oszacowaniem współczynnika regresji,
- liczba osób w gospodarstwie domowym charakteryzowała się różnymi znakami: w modelu popytu potencjalnego – ujemnym, a w modelu popytu zrealizowanego – dodatnim,
- parametry stojące przy zmiennej oszcz3 (w wysokości pomiędzy miesięcznymi a rocznymi dochodami rodziny) w obydwu modelach były dodatnie,
- parametry stojące przy zmiennych reprezentujących kategorie A4 (miasto o liczbie ludności 201-500 tys.) i A5 (miasto o liczbie ludności powyżej 500 tys.) w obydwu modelach były ujemne,
- oszacowania parametrów stojących przy zmiennych reprezentujących kategorie W2 (zasadnicze zawodowe) i W4 (wyższe) w obydwu modelach były ujemne.

Okazało się również, że zarówno w modelach ubezpieczeń na życie, jak i w modelach ubezpieczeń na dożycie dopasowanie modeli w przypadku popytu potencjalnego było lepsze.

Reasumując, należy stwierdzić, że w stosunku do kategorii bazowych przyjętych w modelach regresji dla zmiennych jakościowych odmienne preferencje w stosunku do gospodarstw domowych ze wsi, w których głowa gospodarstwa domowego miała wykształcenie co najwyżej gimnazjalne, nieposiadających oszczędności, miały:

- w zakresie popytu potencjalnego i zrealizowanego na ubezpieczenia na życie gospodarstwa domowe z małych i dużych miast (mniejsze zainteresowanie ubezpieczeniami – ujemne oceny parametrów przy tych zmiennych), w których głowa gospodarstwa domowego miała wykształcenie średnie (większe zainteresowanie ubezpieczeniami – dodatnie oceny parametrów), posiadające oszczędności

w wysokości pomiędzy miesięcznymi a rocznymi dochodami rodziny (większe zainteresowanie ubezpieczeniami – dodatnie oceny parametrów),

- w zakresie popytu potencjalnego i zrealizowanego na dożycie gospodarstwa domowe z dużych i bardzo dużych miast (mniejsze zainteresowanie ubezpieczeniami – ujemne oceny parametrów przy tych zmiennych), w których głowa gospodarstwa domowego miała wykształcenie zawodowe lub wyższe (mniejsze zainteresowanie ubezpieczeniami – ujemne oceny parametrów przy tych zmiennych), posiadające oszczędności w wysokości pomiędzy miesięcznymi a rocznymi dochodami rodziny (większe zainteresowanie ubezpieczeniami – dodatnie oceny parametrów).

## Literatura

- Batóg B., Mojsiewicz M., Wawrzyniak K., *Efektywność metod statystycznej analizy wielowymiarowej jako narzędzia segmentacji rynku ubezpieczeniowego*, [w:] *Klasyfikacja i analiza danych – teoria i zastosowania*, Taksonomia 14, Prace Naukowe Akademii Ekonomicznej we Wrocławiu, AE, Wrocław 2007.
- Gatnar E., Walesiak M., *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*, AE, Wrocław 2004.
- Johnston J., *Econometric Methods*, McGraw-Hill Book Company, 1991.
- Maddala G.S., *Ekonometria*, Wydawnictwo Naukowe PWN, Warszawa 2006.
- Mojsiewicz M., *O pomiarze użyteczności ubezpieczeń*, [w:] *Ubezpieczenia wobec wyzwań XXI wieku*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 1197, AE, Wrocław 2008.

### THE CLASSIFICATION OF HOUSEHOLDS ACCORDING TO POTENTIAL AND ACTUAL DEMAND IN THE INSURANCE MARKET IN POLAND

**Summary:** In the paper, the authors deal with the problems of potential and actual demand for property, life, pure endowment, and health insurances. The classification of households was conducted by means of linear regression models with changing parameters. The declared monthly premium and its transformation were dependent variables in models of potential demand, and number of policies was dependent variable in models of actual demand. The homogeneous groups were identified on the base of the categories of qualitative variables.