

## ROZKŁADY WYDATKÓW MIESZKAŃCÓW OBSZARÓW MIEJSKICH I WIEJSKICH W POLSCE

**Anna Turczak**

Wydział Ekonomii i Informatyki  
Zachodniopomorska Szkoła Biznesu w Szczecinie  
e-mail: aturczak@zpsb.pl

**Patrycja Zwiech (ORCID: 0000-0002-4053-8177)**

Wydział Nauk Ekonomicznych i Zarządzania  
Uniwersytet Szczeciński  
e-mail: patrycjazwiech@tlen.pl

**Streszczenie:** W artykule określono stopień podobieństwa rozkładów wydatków na osobę mieszkańców miast i wsi. Podobieństwo rozkładów zmierzono za pomocą zaproponowanej statystyki  $\lambda$  (lambda), przy czym dwa rozkłady były tym bardziej podobne, czym  $\lambda$  miała mniejszą wartość. Przeprowadzenie procedury klasyfikacji pozwoliło na utworzenie następujących trzech grup: A – miasta o liczbie mieszkańców 500 tys. i więcej, B – miasta o liczbie mieszkańców nie przekraczającej 500 tys., C – obszary wiejskie. W celu wyciągnięcia wniosków końcowych przeanalizowano rozrzut wydatków między grupami A, B i C oraz wewnątrz tych grup.

**Słowa kluczowe:** wydatki na osobę, rozkład, wariancja wewnątrzgrupowa, wariancja międzygrupowa, Polska

### WPROWADZENIE

Wartość wydatków<sup>1</sup> na osobę jest ważną zmienną różnicującą jakość życia Polaków. Celem niniejszego artykułu jest odpowiedź na pytanie, w jakim stopniu

---

<sup>1</sup> Za Głównym Urzędem Statystycznym przyjęto, że na całkowite wydatki składają się wydatki na towary i usługi konsumpcyjne oraz pozostałe wydatki. Wydatki na towary i usługi konsumpcyjne przeznaczone są na zaspokojenie potrzeb gospodarstwa domowego i obejmują towary zakupione za gotówkę (w tym przy użyciu karty płatniczej bądź

klasy miejscowości zamieszkania w Polsce (w skrócie – KMZ) różnią się między sobą rozkładem miesięcznych wydatków na mieszkańca. Osiągnięciu tego celu służyć będzie realizacja następujących zadań badawczych:

1. podzielenie klas miejscowości zamieszkania na grupy o podobnym rozkładzie badanej zmiennej, tj.:
  - a. wyznaczenie wartości miary stanowiącej podstawę podziału KMZ,
  - b. a następnie wyodrębnienie zbiorów KLM jak najbardziej jednorodnych pod względem rozkładu wydatków na mieszkańca;
2. porównanie średnich wydatków na osobę oraz dyspersji tych wydatków w trzech rozpatrywanych zbiorach;
3. ustalenie wielkości zróżnicowania międzygrupowego i wewnątrzgrupowego.

Niniejszy artykuł ma charakter badawczy. Wszystkie zawarte w nim obliczenia przeprowadzono na podstawie nieidentyfikowalnych danych jednostkowych z badania budżetów gospodarstw domowych<sup>2</sup> zrealizowanego przez Główny Urząd Statystyczny<sup>3</sup>. Co warto podkreślić, badanie budżetów gospodarstw domowych przez GUS prowadzone jest metodą reprezentacyjną, która daje możliwość uogólnienia uzyskanych wyników na wszystkie gospodarstwa domowe w Polsce [Budżety... 2015].

Dla każdego gospodarstwa domowego ankietowanego przez Główny Urząd Statystyczny w ramach badania budżetów gospodarstw domowych za 2012 rok wyznaczono wartość wydatków przypadającą na osobę. Informacje zawarte we wspomnianej bazie danych pozwoliły także na przyporządkowanie poszczególnych gospodarstw do odpowiednich klas miejscowości zamieszkania. W tabeli 1 zebrano informacje na temat struktury próby objętej analizą.

---

kredytowej), na kredyt, otrzymane bezpłatnie, jak również spożycie naturalne (tj. towary i usługi konsumpcyjne pobrane na potrzeby gospodarstwa domowego z działalności rolniczej albo działalności gospodarczej na własny rachunek). Natomiast pozostałe wydatki składają się między innymi z: 1) darów przekazanych innym gospodarstwom domowym i instytucjom niekomercyjnym, 2) niektórych podatków (w tym podatku od spadków i darowizn, podatku od nieruchomości, opłaty za wieczyste użytkowanie gruntu), 3) zaliczek na podatek od dochodów osobistych oraz składek na ubezpieczenia społeczne płaconych samodzielnie przez podatnika [Budżety... 2015].

<sup>2</sup> Trzeba jednak zaznaczyć, że za jednostkę statystyczną w niniejszym artykule przyjęto osobę, a nie gospodarstwo domowe.

<sup>3</sup> Bazę za 2012 rok udostępnił GUS na podstawie Umowy nr 20/Z/DI-6-611/632/2013/RM między GUS i Uniwersytetem Szczecińskim.

Tabela 1. Struktura badanej próby

Wyszczególnienie	Gospodarstwa domowe		Osoby	
	liczba	odsetek	liczba	odsetek
Miasta razem, w tym o liczbie mieszkańców:	21685	57,9%	54125	51,4%
≥ 500 tys.	4768	12,7%	10700	10,2%
⟨200 tys., 500 tys.)	3455	9,2%	8396	8,0%
⟨100 tys., 200 tys.)	2719	7,3%	6887	6,5%
⟨20 tys., 100 tys.)	6447	17,2%	16733	15,9%
< 20 tys.	4296	11,5%	11409	10,8%
Wsie	15742	42,1%	51202	48,6%
Ogółem	37427	100,0%	105327	100,0%

Źródło: obliczenia własne na podstawie nieidentyfikowalnych danych jednostkowych z badania budżetów gospodarstw domowych GUS

Populację generalną stanowią wszyscy mieszkańcy Polski. Liczebność tej populacji wyniosła<sup>4</sup> 38534 tys. osób, w tym 23360 tys. to ludność zamieszkała na obszarach miejskich, a 15174 to ludność zamieszkała na obszarach wiejskich. Oznacza to, że badana próba objęła 0,273% zbiorowości generalnej<sup>5</sup>.

Wydatki przypadające na osobę wyznaczono jako iloraz całkowitych wydatków gospodarstwa domowego i liczby jego członków. Założono zatem, że całkowite wydatki gospodarstwa domowego są dzielone po równo na wszystkich jego członków. Przyjęcie takiego upraszczającego założenia było możliwe, bowiem nie stało w sprzeczności z celami niniejszego opracowania, jak również nie miało wpływu na ostateczne wnioski wynikające z przeprowadzonych analiz.

#### PODZIELENIE KLAS MIEJSCOWOŚCI ZAMIESZKANIA NA GRUPY PODOBNE POD WZGLĘDEM ROZKŁADU WYDATKÓW NA OSOBĘ

Rozpatrywaną zmienną oznaczono przez  $X$ . Niech pierwsza próba liczy  $n_1$  elementów, a z druga –  $n_2$  elementów. Przez  $F_{n_1}(x_i)$  i  $F_{n_2}(x_i)$  oznaczone zostaną dystrybuanty empiryczne dotyczące odpowiednio pierwszej i drugiej próby. Przedmiotem analizy są wielkości różnic między wartościami tych dystrybuant. W celu określenia poszczególnych różnic wszystkie obserwacje występujące w tych próbach uporządkowano w kolejności niemalejącej. Następnie dla każdej  $i$ -tej obserwacji obliczono wartości obu dystrybuant empirycznych odpowiednio według wzorów [Witkowski (red.) 2010]:

<sup>4</sup> Podano stan na 30 czerwca 2012 roku.

<sup>5</sup> W przypadku ludności miast było to 0,232%, a w przypadku ludności wsi – 0,337%.

$$F_{n_1}(x_i) = \frac{n_{1sk.}(x_i)}{n_1}, F_{n_2}(x_i) = \frac{n_{2sk.}(x_i)}{n_2} \quad (1)$$

gdzie  $n_{1sk.}(x_i)$  i  $n_{2sk.}(x_i)$  są liczebnościami skumulowanymi.

W następnym kroku dla każdej wartości zmiennej  $X$  obliczono różnicę między dystrybuantami i odszukano największą wartość bezwzględną różnicy między  $F_{n_1}(x_i)$  i  $F_{n_2}(x_i)$ . Znaną wartość oznaczono przez  $D_{1,2}$ . Tak więc miara  $D_{1,2}$  została zdefiniowana jako [Taylor, Emerson 2011]:

$$D_{1,2} = \max_{x_i} |F_{n_1}(x_i) - F_{n_2}(x_i)|. \quad (2)$$

W oparciu o statystykę  $D$  wyznaczono statystykę  $\lambda$  wyrażoną wzorem:

$$\lambda_{1,2} = D_{1,2} \sqrt{n_{1,2}} \quad (3)$$

gdzie [Rószkiewicz 2012]:

$$n_{1,2} = \frac{n_1 n_2}{n_1 + n_2}. \quad (4)$$

W tym kontekście warto wspomnieć, że statystykę  $\lambda$  można byłoby również wykorzystać do przeprowadzenia testu Kołmogorowa-Smirnowa, a celem realizacji tego nieparametrycznego testu istotności byłaby weryfikacja hipotezy, że dwie populacje mają ten sam rozkład.

W tabelach 2 i 3 podano uzyskane wartości odpowiednio statystyki  $D$  i  $\lambda$ . W celu uproszczenia zapisu dla poszczególnych klas miejscowości zamieszkania wprowadzono następujące symbole:

M\_1: miasta o liczbie mieszkańców  $\geq 500$  tys.,

M\_2: miasta o liczbie mieszkańców  $\langle 200$  tys.,  $500$  tys.),

M\_3: miasta o liczbie mieszkańców  $\langle 100$  tys.,  $200$  tys.),

M\_4: miasta o liczbie mieszkańców  $\langle 20$  tys.,  $100$  tys.),

M\_5: miasta o liczbie mieszkańców  $< 20$  tys.,

W: wsie.

Tabela 2. Wartości statystyki  $D$  obliczone dla poszczególnych par KMZ

$D$	M_1	M_2	M_3	M_4	M_5	W
M_1	0,000	0,184	0,210	0,254	0,289	0,396
M_2	0,184	0,000	0,045	0,091	0,125	0,233
M_3	0,210	0,045	0,000	0,050	0,085	0,198
M_4	0,254	0,091	0,050	0,000	0,044	0,157
M_5	0,289	0,125	0,085	0,044	0,000	0,120
M_6	0,396	0,233	0,198	0,157	0,120	0,000

Źródło: jak w tabeli 1

Tabela 3. Wartości statystyki  $\lambda$  obliczone dla poszczególnych par KMZ

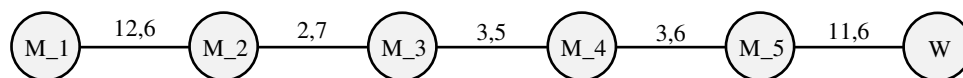
$\lambda$	M_1	M_2	M_3	M_4	M_5	W
M_1	0,0	<b>12,6</b>	13,6	20,5	21,5	37,2
M_2	12,6	0,0	<b>2,7</b>	6,8	8,7	19,8
M_3	13,6	<b>2,7</b>	0,0	3,5	5,6	15,4
M_4	20,5	6,8	<b>3,5</b>	0,0	3,6	17,7
M_5	21,5	8,7	5,6	<b>3,6</b>	0,0	11,6
M_6	37,2	19,8	15,4	17,7	<b>11,6</b>	0,0

Źródło: jak w tabeli 1

Postawionym zadaniem jest podział zbioru sześciu klas miejscowości zamieszkania na takie rozłączne i niepuste podzbiory, aby KMZ należące do tych samych grup były pod względem rozkładów jak najbardziej podobne, a klasy należące do różnych grup były jak najmniej podobne. Klasyfikacja przeprowadzona zostanie w oparciu o miarę  $\lambda$ .

W każdym wierszu tabeli 3 pogrubioną czcionką zaznaczono najmniejsze dodatnie wartości  $\lambda$ . W oparciu o te liczby sporządzono rysunek 1, na którym klasy miejscowości zamieszkania oznaczono kółkami.

Rysunek 1. Graf spójny



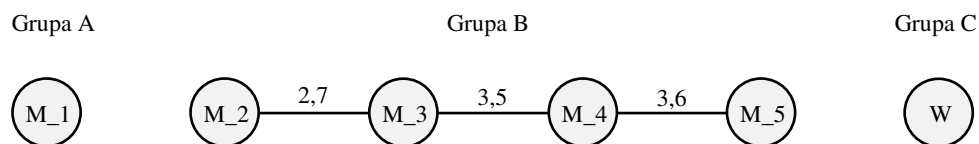
Źródło: opracowanie własne na podstawie tabeli 3

W przypadku potrzeby pogrupowania KMZ w dwa jednorodne podzbiory, z grafu spójnego należałoby usunąć najdłuższe wiązanie odpowiadające odległości równej 12,6. Wtedy w jednej grupie znalazłyby się miasta o liczbie mieszkańców wynoszącej co najmniej 500 tys. (M\_1), a w drugiej grupie – pozostałe pięć KMZ. Postawionym celem jest natomiast wyznaczenie jak najbardziej jednorodnych grup, stąd usunięciu ulega jeszcze wiązanie dla dystansu  $\lambda$  opiewającego na 11,6. W ten sposób powstała następna grupa obejmująca wsie (W). Pozostałe wiązania, które nie zostały usunięte z grafu spójnego, odnoszą się do odległości  $\lambda$  wynoszącej od 2,7 do 3,6 i wartości te można uznać za małe w stosunku do pozostałych (czyli 11,6 i 12,6). Procedurę dzielenia grafu zakończono zatem na etapie wyodrębnienia trzech względnie homogenicznych grup. Są to następujące zbiory:

- zbiór A zawierający element M\_1 (czyli miasta o liczbie mieszkańców wynoszącej co najmniej 500 tys.),
- zbiór B zawierający elementy M\_2, M\_3, M\_4 i M\_5 (czyli miasta o liczbie mieszkańców nie przekraczającej 500 tys.),
- zbiór C zawierający element W (czyli wsie).

Utworzone grupy zaprezentowano na rysunku 2.

Rysunek 2. Podział KMZ na trzy względnie jednorodne grupy



Źródło: opracowanie własne na podstawie rysunku 1

## METODY WYKORZYSTYWANE PRZEZ INNYCH BADACZY PROBLEMU NIERÓWNOŚCI SPOŁECZNYCH

Liczba prac poświęconych analizie wydatków realizowanych przez różne grupy gospodarstw domowych jest dość pokaźna. Niestety w zdecydowanej większości przypadków analiza taka ogranicza się do obliczenia najprostszych charakterystyk opisowych. Do celów przeprowadzenia klasyfikacji rozkładów aparat statystyki opisowej jest jednak zdecydowanie narzędziem niewystarczającym [Szulc 2007].

Autorki niniejszego artykułu przeanalizowały dorobek szeregu badaczy zajmujących się teorią klasyfikacji [Florek i in. 1951, Chomątowski i Sokołowski 1978, Bartosiewicz 1984, Strahl 1987, Brandt 1999, Walesiak 2004, Decker i Lenz (red.) 2007, Jajuga i Walesiak (red.) 2007] oraz aplikacyjnością tych teorii w analizie zjawisk i procesów społeczno-ekonomicznych [Hellwig 1968, Pluta 1977, Podolec i Zając 1978, Wydymus 1984, Pociecha 1986, Pociecha i in. 1988, Nowak 1989, Kruszka 1989, Malina i Zeliaś 1998, Grabiński i in. 1989, Zeliaś (red.) 2000, Zajączkowski 2004, Słaby 2006]. Te szerokie badania literaturowe przeprowadziły w celu znalezienia właściwej metryki pozwalającej na określenie podobieństwa rozkładów, a następnie – w oparciu o wartości tej metryki – dokonanie grupowania rozkładów.

Faktem jest, że w teorii ekonomii istnieje wiele mierników pozwalających na ocenę rozmiarów nierówności występujących między wyodrębnionymi grupami gospodarstw domowych. Mierniki te nie opisują jednak wszystkich różnic w rozkładach porównywanych grup, a bazują jedynie na jakiejś jednej syntetycznej wielkości stanowiącej podstawę porównań. Postawionym w artykule zadaniem jest natomiast wykrycie wszystkich rodzajów występujących różnic – czyli dotyczących rodzaju rozkładu oraz wartości jego poszczególnych parametrów. Jeśli więc obserwacje w dwóch porównywanych ze sobą zbiorowościach mają inny typ rozkładu, to nawet przy tych samych wartościach średnich powinno się je uznać za rozkłady niepodobne. Tak samo jeśli dwie porównywane ze sobą zbiorowości mają taki sam typ rozkładu i taką samą średnią, ale posiadają inne wartości pozostałych parametrów (np. zróżnicowania i/lub asymetrii), to nie mogą być uznane za rozkłady podobne.

## PORÓWNANIE ROZKŁADU ZMIENNEJ W GRUPACH A, B i C

Punktem wyjścia w porównaniu rozkładu miesięcznych wydatków na osobę w trzech grupach zbudowanych w ramach pierwszego zadania badawczego będzie analiza tendencji centralnej, która przeprowadzona zostanie za pomocą średniej arytmetycznej. Wykonane obliczenia doprowadziły do następujących wyników: grupa A – 1595,14 zł/os., grupa B – 1104,80 zł/os., grupa C – 894,43 zł/os. W oparciu o uzyskane liczby można orzec, że średnie wydatki na osobę w miastach o liczbie mieszkańców wynoszącej 500 tys. i więcej były w 2012 roku o 44% wyższe niż w miastach o liczbie mieszkańców nie przekraczającej 500 tys. oraz o 78% wyższe niż na wsiach.

Oczywiście wartość średnia nie wyczerpuje informacji o rozkładzie badanej cechy w zbiorowości [Roeske-Słomka 2010]. Stąd w celu lepszego poznania struktury badanego zjawiska przeanalizowana zostanie również dyspersja.

Dla każdej z grup obliczono odchylenie standardowe. Następnie na podstawie wartości średniej arytmetycznej i odchylenia standardowego wyznaczono współczynnik zmienności. Znajomość średniej i odchylenia standardowego pozwoliła również ustalić tzw. typowy obszar zmienności. Wartości odchyleń standardowych, współczynników zmienności oraz dolnych i górnych granic typowych obszarów zmienności otrzymanych dla grup A, B, C zaprezentowano w tabeli 4.

Tabela 4. Odchylenia standardowe, współczynniki zmienności oraz granice typowych obszarów zmienności dla poszczególnych grup

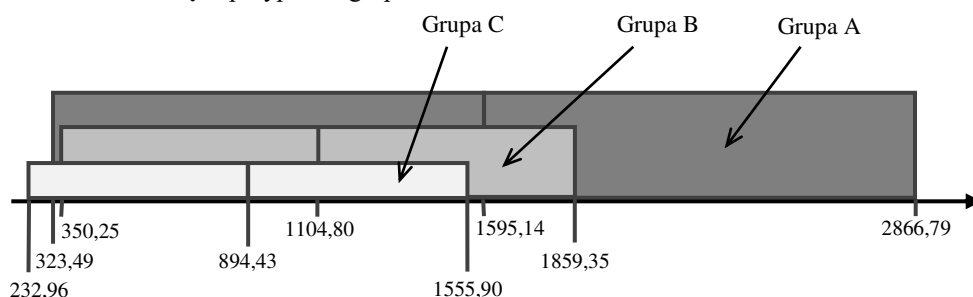
Wyszczególnienie	Grupa A	Grupa B	Grupa C
Odchylenie standardowe (w zł)	1271,65	754,55	661,47
Współczynnik zmienności (w %)	79,7	68,3	74,0
Dolna granica typowego obszaru zmienności (w zł)	323,49	350,25	232,96
Górna granica typowego obszaru zmienności (w zł)	2866,79	1859,35	1555,90
Odsetek mieszkańców, którzy znaleźli się w typowym obszarze zmienności (w %)	90,0	85,9	89,0

Źródło: jak w tabeli 1

Dwóch mieszkańców z grupy A w 2012 roku różniło się od siebie pod względem miesięcznych wydatków przeciętnie o 1271,65 zł, z grupy B – o 754,55 zł, natomiast z grupy C – o 661,47 zł. Analiza danych zamieszczonych w tabeli 4 pozwala stwierdzić, że wszystkie trzy grupy gospodarstw domowych miały podobny poziom zróżnicowania, o czym świadczą zbliżone wartości klasycznego współczynnika zmienności. Jednak największą zmienność można zaobserwować w przypadku grupy A (79,7%), a najmniejszą – w przypadku grupy B (68,3%).

Na rysunku 3 na osi miesięcznych wydatków zobrazowano położenie względem siebie poszczególnych średnich. Pokazano również jak duża jest rozpiętość typowych obszarów zmienności w przypadku każdej z utworzonych grup.

Rysunek 3. Wartości średnie oraz granice typowych obszarów zmienności dla wydatków na osobę w przypadku grup A, B i C



Źródło: opracowanie własne na podstawie tabeli 4

### PORÓWNANIE ROZPROSZENIA WYDATKÓW WEWNĄTRZ GRUP Z ROZPROSZENIEM MIĘDZY GRUPAMI

Trzecim i zarazem ostatnim postawionym zadaniem badawczym jest analiza zróżnicowania międzygrupowego, dzięki której określone zostanie, jak duże są przeciętne różnice między obserwacjami pochodzącymi z poszczególnych grup. Aby oszacować wielkość tego rozrzutu, obliczona będzie wariancja międzygrupowa, odchylenie standardowe międzygrupowe oraz współczynnik zmienności międzygrupowy. W ramach przeprowadzanej analizy określone zostaną również udziały wariancji międzygrupowej i wewnątrzgrupowej w wariancji ogółem. Wariancja ma bowiem pewną własność istotną z punktu widzenia celu niniejszego artykułu. Otóż suma wariancji międzygrupowej i średniej wariancji wewnątrzgrupowej jest zawsze równa wariancji ogólnej obliczonej dla całej badanej zbiorowości statystycznej [Sobczyk 2010].

Wszystkie wyniki uzyskane w ramach realizacji trzeciego zadania badawczego zebrano w tabeli 5.

Tabela 5. Porównanie zróżnicowania międzygrupowego i wewnątrzgrupowego

Wyszczególnienie	Wartość międzygrupowa	Średnia wartość wewnątrzgrupowa	Wartość ogółem (tj. dla trzech grup łącznie)
Wariancja (w zł <sup>2</sup> )	43187,50	611711,94	654899,44
Odchylenie standardowe (w zł)	207,82	782,12	809,26
Współczynnik zmienności (w %)	19,7	74,3	76,9
Struktura wariancji ogółem (w %)	6,6	93,4	100

Źródło: jak w tabeli 1



Porównanie wartości wariancji międzygrupowej ze średnią wartością wariancji wewnątrzgrupowej pozwala orzec, że rozrzut wyników wewnątrz grup znacznie przewyższa rozrzut wyników między grupami. Do tego samego wniosku można dojść na podstawie porównania odpowiednich odchyłeń standardowych, jak również na podstawie porównania odpowiednich współczynników zmienności. Warto przy tym zauważyć, że średnie odchylenie standardowe wewnątrzgrupowe jest prawie czterokrotnie wyższe od odchylenia międzygrupowego. Tak więc przeciętne różnice w wartości wydatków na osobę między dwoma gospodarstwami z tej samej grupy są prawie cztery razy większe niż przeciętne różnice między dwoma gospodarstwami mającymi wydatki *per capita* na poziomie średnim w poszczególnych grupach.

Dodatkową informację na temat znaczenia wariancji międzygrupowej i wariancji wewnątrzgrupowej dla kształtowania się wariancji ogółem dają liczby umieszczone w ostatnim wierszu tabeli 5. Wykazano, że średnia wariancja wewnątrzgrupowa stanowi ponad 93% wariancji całkowitej, a udział wariancji międzygrupowej to pozostałe 7%. Toteż nie ulega wątpliwości, że przeciętne rozbieżności między poziomem wydatków na osobę gospodarstw domowych należących do tej samej grupy należy uznać za bardzo duże na tle takich rozbieżności między średnimi dla gospodarstw domowych z różnych grup.

Podsumowując powyższy wywód trzeba jeszcze raz podkreślić, iż to, czy gospodarstwo domowe znajduje się w mieście o liczbie mieszkańców wynoszącej co najmniej 500 tys., czy w mieście o liczbie mieszkańców poniżej 500 tys., czy na wsi, ma wyraźne znaczenie dla poziomu wydatków na osobę, o czym mogą świadczyć zaistniałe różnice w średnich obliczonych dla każdej z trzech grup wyróżnionych w badaniu. Różnice między tymi średnimi należy jednak uznać za małe w zestawieniu z przeciętnymi odchyleniami, jakie występują między gospodarstwami domowymi wewnątrz poszczególnych grup.

## PODSUMOWANIE

Przedmiotem badania w niniejszym artykule było podobieństwo rozkładów wydatków na osobę w poszczególnych klasach miejscowości zamieszkania w Polsce. Klasy te pogrupowano w trzy zbiory o najbardziej zbliżonych do siebie rozkładach. Grupowania tego dokonano w oparciu o wartości statystyki  $\lambda$ . W efekcie zastosowanej procedury klasyfikacyjnej uzyskano trzy grupy: A – z miastami o liczbie mieszkańców wynoszącej co najmniej 500 tys., B – z miastami o liczbie mieszkańców wynoszącej mniej niż 500 tys., C – z obszarami wiejskimi. Okazało się więc, że polskie miasta o liczbie mieszkańców wynoszącej co najmniej 500 tys. oraz polskie wsie mają na tyle różne rozkłady, iż nie można ich uznać za podobne do którejkolwiek z pozostałych czterech klas miejscowości

zamieszkania. Z kolei miasta o liczbie mieszkańców:

< 20 tys.,

(20 tys., 100 tys.),

(100 tys., 200 tys.),

(200 tys., 500 tys.)

były pod względem rozkładu miesięcznych wydatków zbliżone do siebie w stopniu wystarczającym, aby umieścić je w jednej grupie.

Realizacja ostatniego zadania w ramach przeprowadzonych w niniejszym artykule badań pozwoliła na ustalenie, że – pod względem wydatków na osobę – gospodarstwa domowe w Polsce charakteryzują się dużo większym zróżnicowaniem wewnątrz wyznaczonych trzech grup niż między tymi grupami. Średnie odchylenie standardowe wewnątrzgrupowe ma bowiem wartość kilkakrotnie większą od odchylenia międzygrupowego, a wariancja międzygrupowa stanowi jedynie kilka procent wariancji ogółem. Najwyraźniej zatem na kształtowanie się wydatków na osobę w gospodarstwie domowym ma wpływ wiele dodatkowych czynników oprócz faktu, że gospodarstwo jest zlokalizowane w dużym mieście, w małym mieście czy na wsi. Niewątpliwie wielkość wydatków uzależniona jest również od tego, w jakim miejscu w Polsce – tj. w jakim województwie, w jakim powiecie i jakiej gminie – znajduje się dane gospodarstwo domowe, do jakiej grupy społeczno-ekonomicznej należy, jaka jest jego wielkość oraz typ biologiczny. Ze względu na obszar zainteresowań naukowych autorek szczególnie ciekawe jest jednak przede wszystkim to, jaka jest zależność między wykształceniem członków gospodarstwa domowego a wielkością i strukturą realizowanych przez nich wydatków. W dostępnej literaturze zostało już bowiem wykazane, iż poziom wykształcenia Polaków rzutuje w istotny sposób na standard ich życia i – co ciekawe – można zauważyć pewne przestrzenne prawidłowości w tym zakresie [Czyżewski, Polcyn 2016]. Dociekania na ten temat stały się celem dalszych badań autorek.

## BIBLIOGRAFIA

- Bartosiewicz S. (1984) *Zmienne syntetyczne w modelowaniu ekonometrycznym*. Akademia Ekonomiczna we Wrocławiu.
- Brandt S. (1999) *Analiza danych*. Wydawnictwo Naukowe PWN, Warszawa.
- Budżety gospodarstw domowych 2014 (2015) Główny Urząd Statystyczny, Warszawa.
- Chomątowski S., Sokołowski A. (1978) Taksonomia struktur. *Przegląd Statystyczny*, XXV (2), 217–226.
- Czyżewski B., Polcyn J. (2016) Education quality and its drivers in rural areas of Poland. *Eastern European Countryside*, 22, 197–227, DOI: 10.1515/eec-2016-0010.
- Decker R., Lenz H. J. (red.) (2007) *Advances in data analysis*. Wyd. Springer, Berlin.
- Florek K., Łukaszewicz J., Perkal J., Steinhaus H., Zubrzycki S. (1951) Taksonomia wrocławska. *Przegląd Antropologiczny*, 17, 193–211.

- Grabiński T., Wydymus S., Zeliaś A. (1989) Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych. PWN, Warszawa.
- Hellwig Z. (1968) Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr. *Przegląd Statystyczny*, 4, 307–327.
- Jajuga K., Walesiak M. (red.) (2007) Taksonomia 14. Klasyfikacja i analiza danych – teoria i zastosowania. Akademia Ekonomiczna im. Oskara Langego we Wrocławiu.
- Kruszka K. (1989) Miary podobieństwa struktury obiektów społeczno-ekonomicznych (studium porównawcze). *Zeszyty Naukowe, Seria I*, 159, Prace Instytutu Cybernetyki Ekonomicznej, Akademia Ekonomiczna w Poznaniu, 48–65.
- Malina A., Zeliaś A. (1998) On building taxonomic measures on living conditions. *Statistics in Transition*, 3 (3), 523–544.
- Nowak E. (1989) Metody taksonomiczne w klasyfikacji obiektów społeczno-ekonomicznych. Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Pluta W. (1977) Wielowymiarowa analiza porównawcza w badaniach ekonomicznych. Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Pociecha J. (1986) Statystyczne metody segmentacji rynku. Akademia Ekonomiczna w Krakowie.
- Pociecha J., Podolec B., Sokołowski A., Zając K. (1988) Metody taksonomiczne w badaniach społeczno-ekonomicznych. PWN, Warszawa.
- Podolec B., Zając K. (1978) Ekonometryczne metody ustalania rejonów konsumpcyjnych. PWE, Warszawa.
- Roeske-Słomka I. (2010) Statystyka opisowa. Wyd. UE w Poznaniu.
- Rószkiewicz M. (2012) Metody ilościowe w badaniach marketingowych. PWN, Warszawa.
- Słaby T. (2006) Konsumpcja. Eseje statystyczne. Wyd. Difin, Warszawa.
- Sobczyk M. (2010) Statystyka opisowa. Wyd. C.H. Beck, Warszawa.
- Strahl D. (1987) Dyskryminacja struktur. Wyd. AE we Wrocławiu.
- Szulc A. (2007) Dochód i konsumpcja. [w:] Panek T. (red.) Statystyka społeczna. PWE, Warszawa.
- Taylor B. A., Emerson J. W. (2011) Nonparametric goodness-of-fit tests for discrete null distributions. *The R Journal*, 3 (2), 34–39.
- Walesiak M. (2004) Problemy decyzyjne w procesie klasyfikacji zbioru obiektów. *Ekonometria*, 13, Prace Naukowe Akademii Ekonomicznej we Wrocławiu, 52–71.
- Witkowski M. (red.) (2010) Statystyka matematyczna w zarządzaniu. Wyd. UE w Poznaniu.
- Wydymus S. (1984) Metody wielowymiarowej analizy rozwoju społeczno-gospodarczego. Wyd. AE w Krakowie.
- Zajączkowski S. (2004) Wykorzystanie analizy skupień w badaniach struktury wydatków polskich gospodarstw domowych w latach 1998–2002. *Studia i Prace Kolegium Zarządzania i Finansów SGH*, 52, 144–163.
- Zeliaś A. (red.) (2000) Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym. Wyd. AE w Krakowie.