

Alina Jędrzejczak

Uniwersytet Łódzki
e-mail: jedrzej@uni.lodz.pl

Jan Kubacki

Urząd Statystyczny w Łodzi
e-mail: j.kubacki@stat.gov.pl

ANALIZA ROZKŁADÓW DOCHODU ROZPORZĄDZALNEGO WEDŁUG WOJEWÓDZTW Z UWZGLĘDNIENIEM CZASU

ANALYSIS OF DISTRIBUTION OF AVAILABLE INCOME BY REGIONS USING TIME DEPENDENCIES

DOI: 10.15611/pn.2017.469.07

JEL Classification: C22, C23, C83, D31, R11

Streszczenie: W pracy przedstawiono aproksymację rozkładu dochodu rozporządzalnego, w tym dochodu rozporządzalnego na osobę, na podstawie danych indywidualnych pochodzących z badania budżetów gospodarstw domowych za lata 2003-2011 z użyciem trójparametrowego rozkładu Daguma. Wyznaczono rozkłady tych zmiennych, stosując pakiet VGAM [Yee, Wild 1996; Yee 2015] dla środowiska R-project [R Core Team 2015]. Oceniono również stopień zgodności dopasowania rozkładów teoretycznych do empirycznych. Przedstawiono wstępne wyniki klasyfikacji wybranych parametrów rozkładu Daguma z użyciem hierarchicznych metod aglomeracyjnych oraz metody k -średnich. Oszacowano parametry nieliniowych funkcji trendu dla parametru skali rozkładu Daguma opisującego zmiany dochodu rozporządzalnego na osobę i porównano ten model z modelem trendu dla średniego dochodu rozporządzalnego na osobę. Uzyskano dość dobrą zgodność obu tych modeli, co potwierdzają wykresy ujmujące zależności nieliniowe w czasie.

Słowa kluczowe: badanie budżetów gospodarstw domowych, analiza rozkładu dochodów, rozkład Daguma, analiza nieliniowa.

Summary: The paper presents the analysis of the distribution of available income, including available income per person based on the Household Budget Survey for the years 2003-2011 using the Dagum distribution. The distributions of these variables were obtained using the package VGAM for R-project environment and individual data. Also the quality of estimated parameters is presented. A preliminary analysis of selected parameters of interest using hierarchical clustering and k -means method is shown. The results of nonlinear estimation models for the Dagum distribution scale parameter of available income is assessed and the model of

the average disposable income per person is compared. Obtained fairly good consistency of both of these models is also confirmed using charts applying nonlinear dependence of time.

Keywords: Household Budget Survey, analysis of income distribution, Dagum distribution, nonlinear analysis.

1. Wstęp

Do analizy rozkładów płac i dochodów wykorzystywane są często rozkłady teoretyczne, które z dostateczną dokładnością opisują empiryczne rozkłady dochodów. Pozwala to na pogłębienie analizy oraz stwarza możliwość prognozowania rozkładów dochodów. Dodatkową zaletą podejścia opartego na aproksymacji rozkładów empirycznych za pomocą teoretycznych jest możliwość wyrównania nieprawidłowości w zbiorach danych pochodzących z próby, wynikających najczęściej z braków odpowiedzi. Do aproksymacji rozkładów dochodów w Polsce stosowane były różne funkcje gęstości, a stopień ich zgodności zmieniał się w czasie.

Ogólną teorię rozkładu logarytmiczno-normalnego sformułował David McAlister [1879]. Rozkład ten został po raz pierwszy zastosowany do badania płac i dochodów przez francuskiego ekonomistę i statystyka R. Gibrata [1931], który jednocześnie sformułował prawo proporcjonalnego efektu, dające stochastyczne podwaliny omawianego rozkładu. W Polsce rozkład logarytmiczno-normalny po raz pierwszy zastosował Wiśniewski [1934]. Z jego badań wynikało, że rozkład ten może być wykorzystywany do badania płac pracowników biurowych. Pałowski [1960] i Vielrose [1960] wykazali przydatność tego rozkładu do aproksymacji płac wszystkich pracowników w gospodarce uspołecznionej. Przydatność rozkładu logarytmiczno-normalnego do badania rozkładów płac w Polsce była także przedmiotem prac m.in. Wąsika [1967], Domańskiego [1973] i Kordosa [1968, 1973].

Własności rozkładu logarytmiczno-normalnego zostały szeroko przedstawione w monografii Aitchisona i Browna [1957]. Własności teoretyczne różnorodnych rozkładów, służących m.in. analizie dochodów, znaleźć można też w monografii Kleibera i Kotza [2003]. Po raz pierwszy rozkład gamma został zastosowany do analizy dochodów przez Marcha [1898], a później przez Amoroso [1925]. Jego zgodność z danymi empirycznymi jest w niektórych przypadkach lepsza niż logarytmiczno-normalnego, a poza tym parametry tego rozkładu można otrzymać w sposób analityczny przy wykorzystaniu metody największej wiarygodności. Rozkłady logarytmiczno-normalny i gamma, przydatne w gospodarce centralnie planowanej, nie znajdują szerszego zastosowania w gospodarce rynkowej, dla której kształtowanie się dochodów można objaśnić z użyciem procesu stochastycznego, którego rozkładem granicznym jest rozkład Daguma. Celem artykułu była próba klasyfikacji województw w Polsce z punktu widzenia parametrów rozkładu dochodów z wykorzystaniem modelu Daguma.

2. Rozkład Daguma jako teoretyczny model rozkładu dochodów

Rozkład Daguma powstał na podstawie pogłębionych obserwacji i analiz empirycznych rozkładów płac i dochodów. Analizy te wykazały, że elastyczność dochodową dystrybuanty można przedstawić w postaci następującego równania różniczkowego [Dagum 1977]:

$$\varepsilon(y, F(y)) = \frac{d \ln F(y)}{d \ln y} = \beta_1 [1 - [F(y)^{\beta_2}]]$$

dla $y \geq 0$, $\beta_1, \beta_2 > 0$.

Z równania powyższego wynika, że elastyczność dochodowa dystrybuanty rozkładu dochodów jest funkcją malejącą, wklęsłą i ograniczoną ($\lim_{F \rightarrow 0^+} \varepsilon(y, F) = \beta\delta$), nie zaś nieograniczoną, jak twierdził Pareto.

Rozwiązując powyższe równanie Dagum otrzymał dystrybuantę rozkładu dochodów:

$$F(y) = \begin{cases} (1 + \lambda y^{-\delta})^{-\beta}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

dla $\beta, \lambda, \delta > 0$,

gdzie: $\beta = 1 / \beta_1$, $\delta = \beta_1 \beta_2$, $\lambda = \exp c$, c – stała całkowania wynikająca z rozwiązania równania różniczkowego.

Momenty zwykłe rzędu r określone są wzorem

$$\mu_r = \beta \lambda^{r/\delta} B(1 - r/\delta, \beta + r/\delta) \quad \text{dla } r < \delta$$

gdzie: $B(\cdot, \cdot)$ oznacza funkcję beta

Parametr λ , pełniący rolę parametru skali rozkładu Daguma, pozwala porównywać rozkłady dochodów dla różnych podpopulacji oraz śledzić zmiany rozkładu danej populacji w czasie. Oprócz parametru skali λ , skorelowanego ze średnim dochodem, również pozostałe parametry rozkładu Daguma mogą być traktowane jako wskaźniki ekonomiczne opisujące różnice między województwami. W odróżnieniu od parametru λ odpowiadają one za nierównomierność rozkładu. Parametry β i δ , będąc parametrami określającymi kształt rozkładu, mają także wpływ na to, czy funkcja określająca gęstość rozkładu Daguma osiąga w jakimś punkcie maksimum. Aby tak było, iloczyn tych parametrów musi być większy od jedności: $\beta\delta > 1$. Jeśli zaś $0 < \beta\delta < 1$, to rozkład pozostaje zeromodalny. Z rozważań tych wynika, że rozkład Daguma może być wykorzystywany zarówno do analizy rozkładów jednomodalnych, jak i zeromodalnych, co bardzo rozszerza zakres jego zastosowań.

Oprócz rozkładu trójparametrowego wyrażonego jak powyżej, który można określić jako rozkład Daguma typu I, Dagum [1977] zaproponował także rozkłady czteroparametrowe, uwzględniające przypadki, kiedy w zbiorowości występują dochody zerowe i ujemne (rozkład typu II); lub też tylko dochody większe od zera (rozkład typu III). Dystrybuanta rozkładu czteroparametrowego może być wyrażona w następującej postaci:

$$F(y) = a + (1 - \alpha)(1 + \lambda y^{-\delta})^{-\beta},$$

dla $\beta, \lambda, \delta > 0$.

Typ rozkładu określonego powyższym wzorem zależy od wielkości parametru α

- $\alpha = 0$ – rozkład typu I ($y > 0, F(0) = 0$),
- $0 < \alpha < 1$ – rozkład typu II ($y > 0, F(0) = \alpha$),
- $\alpha < 0$ – rozkład typu III ($y > y_0 > 0, F(y_0) = 0$).

3. Wykorzystanie rozkładu Daguma do klasyfikacji rozkładów dochodów w Polsce

Otrzymane w pracy wyniki uzyskano na podstawie danych jednostkowych pochodzących z Badania Budżetów Gospodarstw Domowych za lata 2003-2011. W trakcie obliczeń zastosowano wagi dla gospodarstw domowych, pochodzące ze zbiorów zawierających wspomniane dane jednostkowe. Więcej informacji na temat konstrukcji tych wag można znaleźć w części „Uwagi metodologiczne” umieszczonej w publikacjach *Budżety gospodarstw domowych* [GUS 2003-2011], jak również w opracowaniu *Metodologia badania budżetów gospodarstw domowych* [GUS 2011]. W opracowaniach tych przedstawiono m.in. metody wykorzystane przy zbieraniu informacji o dochodach gospodarstw domowych oraz definicje stosowanych w Badaniu Budżetów zmiennych opisujących dochód.

W tabeli 1 przedstawiono wartość średniego miesięcznego dochodu rozporządzalnego na osobę, którą wyznaczono korzystając z estymatora ilorazu (por. wzory (2) i (3) opracowania *Metodologia badania budżetów gospodarstw domowych*), zaś ocena błędu estymacji dla tej wielkości została wyznaczona jako względny błąd estymacji (*Relative Estimation Error* – REE), który określono przez podzielenie standardowego błędu szacunku otrzymanego z użyciem metody zrównoważonych próbek replikacyjnych (BRR) przez ocenę wartości średniej estymatora. W tabeli przedstawiono ponadto ocenę skali dla rozkładu dochodu rozporządzalnego na osobę. Ocena parametru skali wynika z następującej relacji między parametrem skali i parametrem λ rozkładu Daguma: (*parametr skali*) ^{δ} = λ . Z kolei wartość REE dla oceny skali rozkładu Daguma, a także sam szacunek parametru skali pochodzi z estymacji metodą IRLS (*Iteratively Reweighted Least Squares*) wykorzystywaną w pakiecie VGAM. Więcej informacji na ten temat znaleźć można w dokumentacji

pakietu VGAM¹, a w szczególności w opisie opcji `dagum` funkcji `vglm` oraz funkcji `summary` stosowanej do opisu wyników otrzymanych z użyciem funkcji `vglm`. Oceny parametrów β i δ wynikają z odpowiadających tym parametrom oznaczeniom dla wzoru na dystrybuantę rozkładu Daguma. Wskaźnik podobieństwa struktur wynika z sumy minimum dwu wielkości:

- 1) różnicy wielkości dystrybuanty teoretycznej dla określonych, sąsiadujących granic histogramu odpowiadającemu rozkładowi empirycznemu,
- 2) względnych częstości empirycznych dla histogramu reprezentującego rozkład empiryczny wyznaczony jak w punkcie poprzednim.

Rezultaty przedstawione w pracy wynikały z obliczeń przeprowadzonych w kilku etapach. W obliczeniach stosowano głównie środowisko R-project. Najpierw wyznaczono parametry rozkładu Daguma z użyciem pakietu VGAM (w wersji 0.7-1), następnie wykonano analizę skupień dla parametrów rozkładu Daguma, stosując pakiet `cluster` [Maechler i in. 2015] (używano funkcji `agnes`) oraz funkcję `kmeans` z pakietu `stats` dla metody k -średnich. Ponadto używano funkcji `nls` w celu wyznaczenia modeli nieliniowych, a także pakietów `sae` [Molina, Marhuenda 2015] i `sae2` [Fay, Diallo 2015], przy pomocy których wyznaczano parametry modeli dla małych obszarów. Wizualizacji dokonano m.in., stosując pakiet `plot3D` (z użyciem funkcji `persp3D`), zaś dane wczytywane były z użyciem pakietu `RODBC`.

Jak wynika z przedstawionych rezultatów aproksymacji, możliwie jest określenie parametrów rozkładu dla dochodu rozporządzalnego na osobę (jak również samego dochodu rozporządzalnego gospodarstwa) z użyciem rozkładu Daguma. Możliwe jest to zarówno dla całej Polski, jak i dla województw. Jest to tym bardziej uzasadnione, gdyż oceny błędów szacunku parametrów skali takiego rozkładu są dość niskie i świadczą o poprawności zastosowanej metody. Widoczne jest to np. w tabeli 1, gdzie porównano wyniki dla parametru skali rozkładu Daguma oraz szacunków dochodu rozporządzalnego na osobę z użyciem oceny bezpośredniej. O zgodności tego rozkładu z obserwacjami empirycznymi świadczą również wartości współczynników podobieństwa struktur, które – dla dochodu rozporządzalnego na osobę – zmieniają się od 0,9697 do 0,9978, przy czym ich średnia równa się 0,9889. Z kolei miara ta dla rozkładów wyznaczonych dla całej Polski zmienia się od 0,9948 do 0,9968, natomiast ich średnia równa się 0,9960. Zgodność ta jest zachowana również dla dochodu rozporządzalnego gospodarstwa i zmienia się ona od 0,9586 do 0,9964, zaś ich średnia wynosi 0,9841, zatem jest nieco niższa niż dla dochodu rozporządzalnego na osobę.

Z przedstawionych powyżej zależności można wywnioskować, że zmiany parametrów skali rozkładu Daguma objaśniają zmiany średniego dochodu rozporządzalnego na osobę, co zostanie zilustrowane w dalszej części artykułu. Ilustracje dotyczące klasyfikacji wybranych parametrów otrzymanych w drodze analizy za-

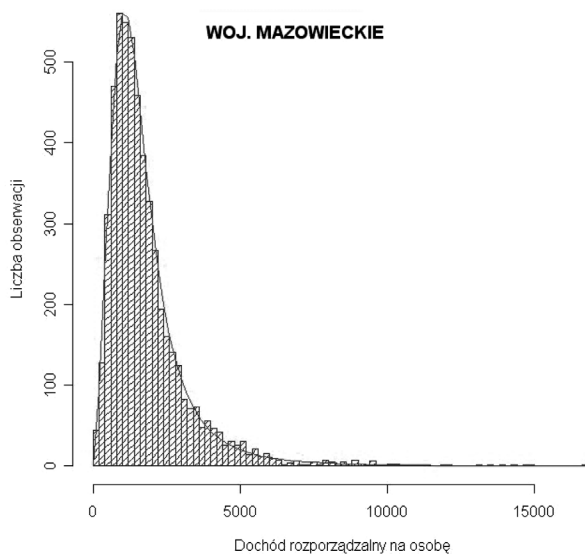
¹ Dokumentacja ta dostępna jest na stronie: <https://cran.r-project.org/web/packages/VGAM/VGAM.pdf> (3.11.2015).

równy samego dochodu rozporządzalnego, jak i parametrów rozkładu Daguma przedstawiono na rys. 3-6. Obrazują one klasyfikację z użyciem metody średniego połączenia odpowiednio dla samego dochodu rozporządzalnego na osobę, parametru skali rozkładu Daguma oraz parametru δ rozkładu Daguma. W przypadku samego dochodu rozporządzalnego widoczny jest dość oczywisty podział wynikający m.in. z poziomu dochodów oraz zróżnicowania regionalnego Polski. Widoczna jest wyróżniająca się pozycja województwa mazowieckiego oraz skupienie dla województw podkarpackiego, lubelskiego, podlaskiego i – niekiedy – świętokrzyskiego. Pozostałe województwa nie tworzą powtarzających się z roku na rok skupień, choć można też wyodrębnić tutaj niekiedy skupienia dla województw z zachodniej części Polski. W przypadku parametru skali klasyfikacja jest już mniej jednoznaczna, co zresztą potwierdza nieco słabsza zależność tego parametru od wartości PKB *per capita* oraz słabsza zależność od czasu w modelu nieliniowym. Niemniej widoczne są tutaj niekiedy wyróżniające się obiekty reprezentujące województwo opolskie lub województwo podkarpackie. Koreponduje to również z danymi liczbowymi przedstawionymi w tabeli 1. Pozostałe województwa nie tworzą wyraźnych skupień, co

Tabela 1. Wyniki estymacji dla dochodu rozporządzalnego na osobę według województw dla roku 2011 wykonane z użyciem estymatora bezpośredniego, estymatora skali rozkładu Daguma, parametrów opisujących kształt rozkładu Daguma oraz wskaźnika podobieństwa struktur

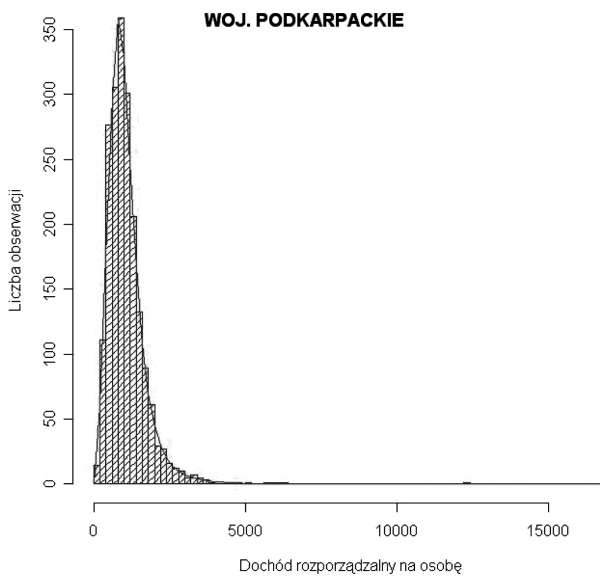
Województwo	Estymator bezpośredni		Ocena skali rozkładu Daguma		Parametr δ	Parametr β	Wsp. podobieństwa struktur
	ocena	REE	ocena	REE	ocena	ocena	ocena
Dolnośląskie	1282,93	2,68	1315,34	0,078	3,319	0,869	98,84
Kujawsko-pomorskie	1108,94	2,17	1216,92	0,085	3,336	0,768	99,23
Lubelskie	1025,80	2,07	1209,24	0,085	3,362	0,633	98,50
Lubuskie	1189,89	1,55	1217,88	0,086	3,674	0,862	98,86
Łódzkie	1203,19	2,62	1326,95	0,078	3,633	0,727	98,10
Małopolskie	1156,79	2,53	1172,92	0,088	3,333	0,923	99,28
Mazowieckie	1622,96	2,02	1522,88	0,068	2,692	0,906	98,10
Opolskie	1181,90	1,88	1452,86	0,072	3,728	0,634	98,34
Podkarpackie	937,85	2,52	1105,42	0,093	3,827	0,680	99,34
Podlaskie	1224,92	1,45	1229,35	0,085	3,344	0,806	99,28
Pomorskie	1286,94	3,09	1506,72	0,069	3,078	0,683	99,10
Śląskie	1215,44	0,95	1400,17	0,073	3,946	0,729	99,42
Świętokrzyskie	1062,78	2,37	1196,73	0,087	3,637	0,733	99,01
Warmińsko-mazurskie	1096,87	2,63	1167,81	0,090	3,283	0,821	98,97
Wielkopolskie	1135,02	2,73	1262,87	0,081	3,350	0,751	99,52
Zachodniopomorskie	1231,10	3,16	1299,15	0,080	3,238	0,842	98,89

Źródło: opracowanie własne.



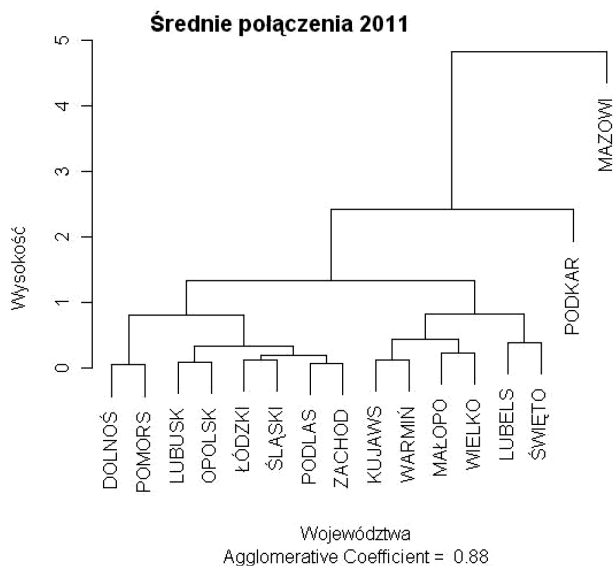
Rys. 1. Rozkład dochodu rozporządzalnego na osobę dla województwa mazowieckiego i roku 2011 – dopasowanie rozkład Daguma (współczynnik podobieństwa struktur – 0,9810)

Źródło: opracowanie własne.



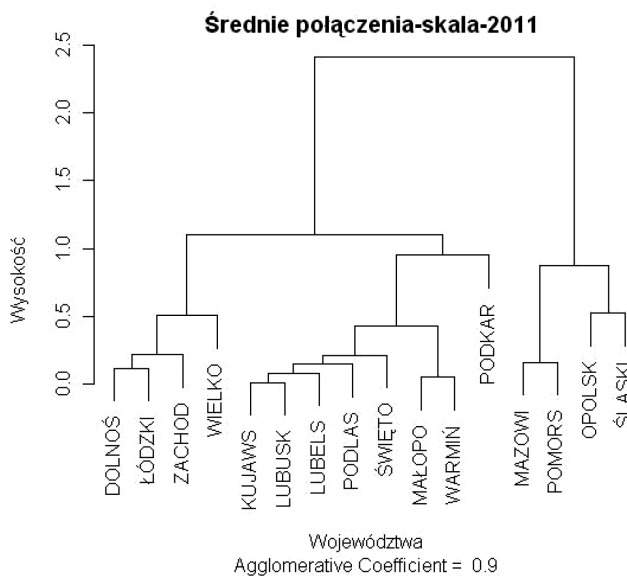
Rys. 2. Rozkład dochodu rozporządzalnego na osobę dla województwa podkarpackiego i roku 2011 – dopasowanie rozkład Daguma (współczynnik podobieństwa struktur – 0,9934)

Źródło: opracowanie własne.



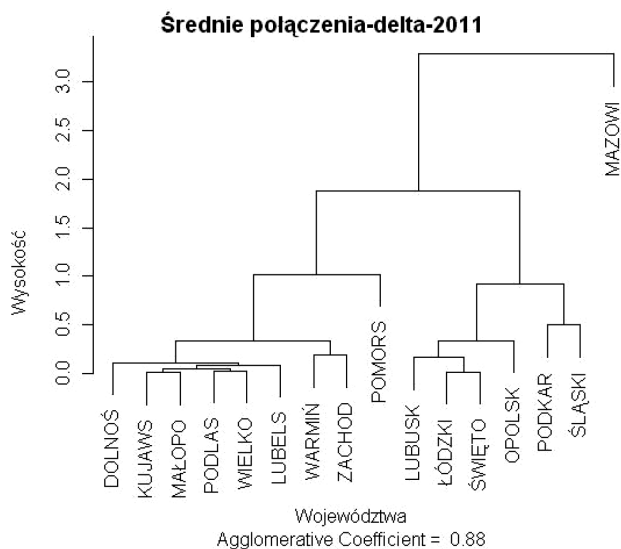
Rys. 3. Dendrogram dla analizy skupień przeprowadzonej metodą średniego wiązania z użyciem odległości euklidesowej dla dochodu rozporządzalnego na osobę według województw dla roku 2011

Źródło: opracowanie własne.



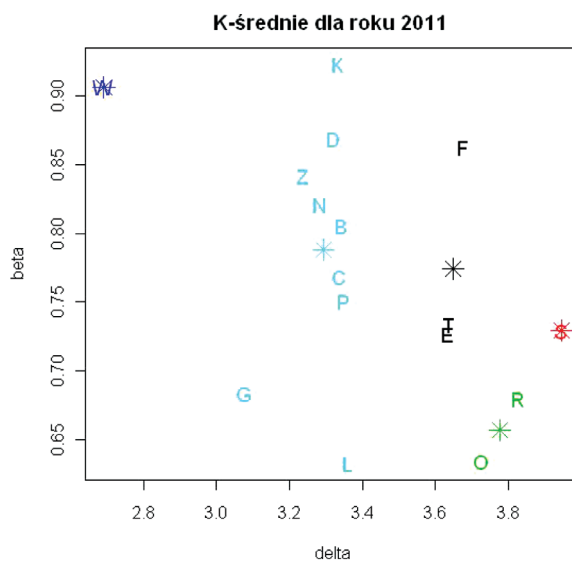
Rys. 4. Dendrogram dla analizy skupień przeprowadzonej metodą średniego wiązania z użyciem odległości euklidesowej dla parametru skali rozkładu Daguma dotyczącego dochodu rozporządzalnego na osobę według województw dla roku 2011

Źródło: opracowanie własne.



Rys. 5. Dendrogram dla analizy skupień przeprowadzonej metodą średniego wiązania z użyciem odległości euklidesowej dla parametru δ rozkładu Daguma dotyczącego dochodu rozporządzalnego na osobę według województw dla roku 2011

Źródło: opracowanie własne.



Rys. 6. Wykres k -średnich dla parametrów β i δ rozkładu Daguma dotyczącego dochodu rozporządzalnego na osobę według województw dla roku 2011 (symbole województw odpowiadają oznaczeniom województw na tablicach rejestracyjnych).

Źródło: opracowanie własne.

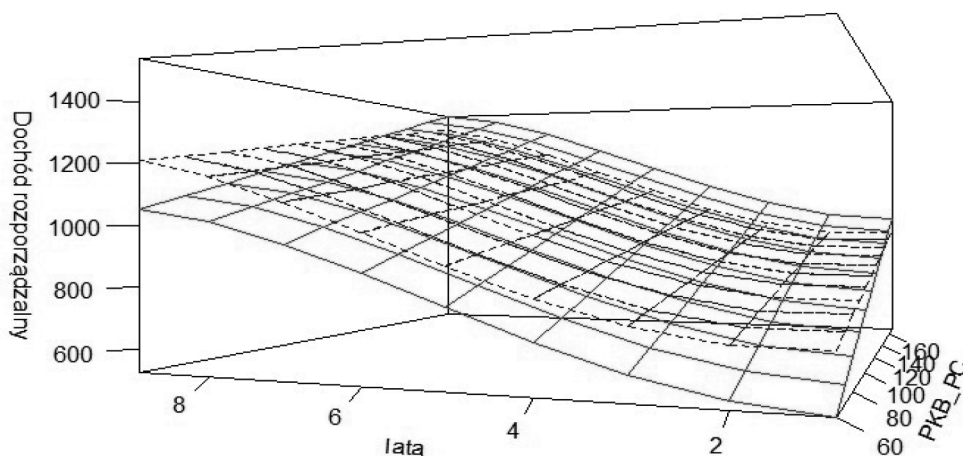
może świadczyć, iż zróżnicowanie tego parametru jest dla pozostałych województw słabsze. Dość ciekawa jest analiza parametru δ , który wskazuje na pewną zależność od poziomu dochodu w danym województwie. Widoczne jest to m.in. dla województwa mazowieckiego, jak również dla województw śląskiego i opolskiego.

Analiza korelacji oraz regresji przeprowadzona dla parametru δ , β i PKB *per capita* wskazała na dość wyraźną ujemną korelację między δ i β oraz istnienie dość wyraźnej ujemnej korelacji między δ i PKB *per capita* oraz słabszej korelacji dodatniej między β i PKB *per capita*. Zależności te widoczne są też dla analizy wykonanej metodą *k*-średnich dla parametrów δ i β . Również tutaj widoczne jest odrębne skupienie dla województwa mazowieckiego oraz wyróżniające się niekiedy skupienia zawierające województwo śląskie i opolskie.

Interesujące może być tutaj porównanie modeli dla dochodu rozporządzanego wykonane z użyciem zależności nieliniowych oraz podobnego modelu dla parametru skali rozkładu Daguma. Modele te mają następującą postać.

$$y_{i,t} = a + bx_{i,t} + ct + dt^2 + et^3,$$

przy czym zmienna t przebiega wartości od 1 do 9 dla lat 2003-2011, zaś $x_{i,t}$ określa wartość PKB *per capita* w województwie i dla czasu t .



Rys. 7. Zależność dochodu rozporządzalnego na osobę dla lat 2003-2011 i województw z użyciem modelu nieliniowego (linia ciągła) oraz nieliniowego modelu dla współczynnika skali rozkładu Daguma (linia przerywana)

Źródło: opracowanie własne.

Ciekawe zależności obserwuje się również, porównując modele nieliniowe obu wspomnianych wielkości. Oceny wartości parametrów tych modeli umieszczono w tabelach 2 i 3. Mają one takie same znaczenie jak w klasycznym modelu regresji. Widoczna tutaj jest wyraźnie inna wartość wyrazu wolnego dla obu tych wielkości.

Jednakże dla pozostałych parametrów zachowany jest ich znak oraz – częściowo – sama wartość, co jest szczególnie widoczne dla parametrów opisujących zależności nieliniowe od czasu, i to zarówno dotyczące składnika w drugiej potęgze, jak i składnika w potęgze trzeciej. Mniejsza jest jednak istotność dla tych dwu parametrów w przypadku modelu dla współczynnika skali rozkładu Daguma. Charakter tych zależności dobrze ilustruje rys. 7.

Tabela 2. Zestawienie cech diagnostycznych modelu dla województw dla dochodu rozporządzalnego na podstawie danych z Badania Budżetów Gospodarstw Domowych dla lat 2003-2011 z użyciem modelu nieliniowego.

Zmienna	Estymator	Błąd stand.	t-Studenta	Wartość <i>p</i>	Istotność
Wartość $R^2 = 0,9667$, wartość skorygowanego $R^2 = 0,9326$					
a	277,06	36,7547	7,5382	5,57E-12	***
b	4,45	0,2197	20,2744	< 2e-16	***
c	-44,88	25,1334	-1,7856	0,07634	.
d	26,22	5,6896	4,6091	9,07E-06	***
e	-1,66	0,3755	-4,4318	1,88E-05	***

Źródło: opracowanie własne.

Tabela 3. Zestawienie cech diagnostycznych modelu dla województw dla parametru skali rozkładu Daguma wyznaczonego dla dochodu rozporządzalnego na podstawie danych z Badania Budżetów Gospodarstw Domowych dla lat 2003-2011 z użyciem modelu nieliniowego

Zmienna	Estymator	Błąd stand.	t-Studenta	Wartość <i>p</i>	Istotność
Wartość $R^2 = 0,9015$, wartość skorygowanego $R^2 = 0,8073$					
a	594,64	61,5708	9,6578	3,36E-17	***
b	2,40	0,3680	6,5092	1,27E-09	***
c	-74,96	42,1031	-1,7803	0,0772	.
d	29,37	9,5312	3,0810	0,0025	**
e	-1,69	0,6290	-2,6863	0,0081	**

Źródło: opracowanie własne.

Zależności nieliniowe mają też podobny charakter w przypadku, gdy analizę rozszerzy się na model oceny przeciętnego dochodu na osobę dla małych obszarów, uwzględniający niepewność szacunków bezpośrednich. Zwraca tutaj uwagę zbliżona wartość ocen parametrów modelu oraz przeciętnie większa efektywność takiego modelu niż szacunków bezpośrednich, a nawet modelu Rao-Yu wykorzystującego zależności czasowe.

4. Zakończenie

Z przedstawionych w pracy rezultatów można wnosić, iż jest możliwe efektywne dopasowanie rozkładu dochodu rozporządzalnego (w tym rozkładu tego dochodu na osobę) z użyciem rozkładu Daguma. Istnieje też możliwość użycia modeli przedstawiających zależność tych wielkości od czasu, mających dość dobre własności statystyczne. Użycie modeli dla małych obszarów potwierdza również zasadność użycia estymacji nieliniowej.

Z analizy skupień przeprowadzonych dla parametrów rozkładu Daguma wynika, iż zróżnicowanie tych parametrów daje się częściowo objaśnić poziomem dochodów w danym województwie, chociaż tendencje te nie są jednoznaczne i zmieniają się z roku na rok.

Wydaje się interesujące, czy podobną analizę można by przedstawić dla innych zmiennych dochodowych. Ponadto zasadne może być przeprowadzenie bardziej szczegółowej analizy parametrów rozkładu Daguma. Wydaje się, iż użycie tutaj metod klasyfikacyjnych dla wielu zmiennych mogłoby być w tym przypadku pomocne. Jednakże ze względu na zakres pracy przedstawienie pełnej analizy tego typu może być utrudnione.

Literatura

- Aitchison J., Brown J.A.C., 1957, *The Lognormal Distribution*, University Press, Cambridge.
- Amoroso L., 1924-1925, *Ricerca intorno alla curva dei redditi*, Annali di Matematica Pura ed Applicata, Series 4-21, s. 123-157.
- Dagum C., 1977, *A new model of personal income distribution. Specification and estimation*, Economie Appliquée, XXX(3), s. 413-436.
- Domański Cz., 1973, *Struktura płac pracowników w województwach grodzkich według działów gospodarki narodowej*, Studia Prawno-Ekonomiczne, nr 11.
- Fay R.E., Diallo M., 2015, *sae2: Small Area Estimation: Time-series Models, package version 0.1-1*, <https://cran.r-project.org/web/packages/sae2/index.html> (1.04.2015).
- Gibrat R., 1931, *Les Inégalités Économiques*, Librairie du Recueil Sirey, Paris.
- Główny Urząd Statystyczny, 2003-2011, *Budżety gospodarstw domowych*, <http://stat.gov.pl/obszary-tematyczne/warunki-zycia/dochody-wydatki-i-warunki-zycia-ludnosci/metodologia-badania-budzetow-gospodarstw-domowych,10,1.html> (24.05.2013).
- Kleiber Ch., Kotz S., 2003, *Statistical Size Distributions in Economics and Actuarial Sciences*, John Wiley & Sons, Inc., Hoboken, New Jersey.
- Kordos J., 1968, *Metody matematyczne badania i analizy rozkładu dochodów ludności*, Główny Urząd Statystyczny, Warszawa.
- Kordos J., 1973, *Metody analizy i prognozowania rozkładów płac i dochodów ludności*, Główny Urząd Statystyczny, Warszawa.
- Maechler M., Rousseeuw P., Struyf A., Hubert M., Hornik K., 2015, *cluster: Cluster Analysis Basics and Extensions*. R package version 2.0.3, <https://cran.r-project.org/web/packages/cluster/index.html> (10.03.2016).

- March L., 1898, *Quelques exemples de distribution des salaires*, Journal de la Société statistique de Paris, 39, s. 193-206 i 241-248.
- McAlister, D., 1879, *The law of the geometric mean*, Proceedings of the Royal Society of London, 29, s. 367-375.
- Molina I., Marhuenda Y., 2015, *sae: An R package for small area estimation*, The R Journal, 7(1), s. 81-98, <http://journal.r-project.org/archive/2015-1/molina-marhuenda.pdf> (29.07.2015).
- Pawłowski Z., 1960, *Badanie elastyczności popytu a rozkład dochodów*, Przegląd Statystyczny, 7, s. 445-465.
- R Core Team, 2015, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org> (10.03.2016).
- Vielrose E., 1960, *Rozkład dochodów według wielkości*, Polskie Wydawnictwo Gospodarcze, Warszawa.
- Wąsik B., 1967, *Dwu- i trójparametryczny rozkład logarytmiczno-normalny jako aproksymacja rozkładów zarobków pracowników gospodarki uspołecznionej 1955-1965*, Przegląd Statystyczny, 14(4), s. 409-424.
- Wiśniewski J., 1934, *Rozkład dochodów według wysokości*, Instytut Badań Koniunktur Gospodarczych I Cen, Warszawa.
- Yee T.W., 2015, *Vector Generalized Linear and Additive Models: With an Implementation in R*, Springer, New York.
- Yee T.W., Wild C.J., 1996, *Vector generalized additive models*, Journal of Royal Statistical Society, Series B, 58(3), s. 481-493.