

*Anna Turczak<sup>1</sup>, Patrycja Zwiech<sup>2</sup>*

<sup>1</sup>Zachodniopomorska Szkoła Biznesu w Szczecinie, e-mail: aturczak@zpsb.pl

<sup>2</sup>Uniwersytet Szczeciński, e-mail: patrycjazwiech@tlen.pl

## Zastosowanie testu Kołmogorowa-Smirnowa do badania zgodności rozkładów dochodu mieszkańców obszarów wiejskich i miejskich w Polsce

**Zarys treści:** Celem artykułu jest odpowiedź na pytanie, czy rozkład dochodu rozporządzalnego na osobę dla mieszkańców polskich miast jest taki sam jak dla mieszkańców polskich wsi, a jeżeli nie, to jakie są występujące w tym względzie różnice. Do weryfikacji odpowiednich hipotez statystycznych posłużył test Kołmogorowa-Smirnowa. Obliczenia przeprowadzono oddzielnie dla piętnastu kolejnych lat. Wyniki badań wskazują, że średni dochód rozporządzalny na osobę jest znacznie większy w miastach niż na wsiach i właśnie różnice w zakresie tendencji centralnej są główną przyczyną faktu, że rozpatrywanych rozkładów nie można uznać za identyczne. Można natomiast przyjąć, że zróżnicowanie dochodu na polskich wsiach i w polskich miastach jest podobne, a asymetria tych rozkładów jest prawie identyczna.

**Słowa kluczowe:** dochód rozporządzalny na osobę, rozkład, test Kołmogorowa-Smirnowa

### Wprowadzenie

Na gruncie nauk społecznych można wyróżnić następujące wymiary nierówności (Wójcik-Żołądek 2013, s. 2):

- wymiar ekonomiczny, który obejmuje takie kategorie, jak dochód, majątek, warunki bytowe;
- wymiar społeczny, który dotyczy wykształcenia, stylu życia, uczestnictwa w kulturze, prestiżu społecznego;
- wymiar polityczny, który odnosi się do różnic w udziale we władzy oraz w zaangażowaniu obywatelskim;
- wymiar przestrzenny, który wynika z miejsca usytuowania i geograficznego zasięgu funkcjonowania.

Traktowanie aspektu ekonomicznego jako jedyne go wymiaru nierówności w społeczeństwie jest więc zbyt n i m u p r o s z c z e n i e m . W l i t e r a t u r z e p r e d m i o t u r o z w a r s t w i e n i e d o c h o d o w e u z n a w a n e j e s t j e d n a k z a j e d e n z n a j i s t o t n i e j s z y c h m i e r n i k ó w n i e r ó w n o ś c i , b o w i e m p o z i o m d o c h o d u u w a ż a n y j e s t p o w s z e c h n i e z a n a j w a ż n i e j s z y w y z n a c z n i k s t a t u s u s o c j e l n e g o . P o d k r e ś l a s i ę r ó w n i e ż , ż e d o c h ó d j e s t c z y n n i k i e m r z u t u j ą c y m n a a k t y w n o ś ć j e d n o s t e k i g o s p o d a r s t w d o m o w y c h w n i e m a l w s z y s t k i c h s f e r a c h ż y c i a – o d k s z t a ł t o w a n i a w a r u n k ó w m a t e r i a l n y c h , p o p r z e z d o s t ę p d o s ł u ż b y z d r o w i a , z a p e w n i e n i e o d p o w i e d n i e g o w y k s z t a ł c e n i a , u c z e s t n i c t w o w k u l t u r z e , d o s t ę p d o z d o b y c z y t e c h n i k i , a ż p o d o s t ę p d o w ł a d z y . T o t e ż n a p o t y k a n i e o g r a n i c z e Ń d o c h o d o w y c h z a w ę ż a p o l e d e c y z y j n e g o s p o d a r s t w a d o m o w e g o n i e t y l k o w z a k r e s i e r o z m i a r ó w r e a l i z o w a n e j k o n s u m p c j i , a l e t ę ż p r e z ą d a o s t o p n i u n i e z a s p o k o j e n i a w i e l u i n n y c h p o t r z e b , w t y m p o z a e k o n o m i c z n y c h ( L e s z c z y Ń s k a 2 0 1 4 , s . 4 1 0 ) . M o ż n a n a w e t p o k u s i ć s i ę o s t w i e r d z e n i e , ż e w i e l k o ś ć d o c h o d u , m a j ą c w p ł y w n a r e a l i z a c j ę s z e r o k i e j w i ą z k i c e l ó w m a t e r i a l n y c h i n i e m a t e r i a l n y c h , j e s t g ł ó w n ą – a l e o c z y w i ś c i e n i e j e d y n ą – d e t e r m i n a n t ą p o c z u c i a z a d o w o l e n i a z o g ó l n e j j a k o ś c i ż y c i a c z ł o w i e k a ( B a l 2 0 1 2 , s . 2 5 2 ) .

P r e d s t a w i c i e l e p o s z c z e g ó l n y c h n u r t ó w w e k o n o m i i p r e z e n t u j ą r ó ż n e , n i e j e d n o k r o t n i e s k r a j n i e o d m i e n n e p o d e j ś c i a d o p r o b l e m u w y s t ę p o w a n i a n i e r ó w n o ś c i d o c h o d o w y c h w s o c j e t e t n i e . R ó ż n i c e w p o d e j ś c i u z a s a d z a j ą s i ę z a z w y c z a j w o s o b i s t y c h p r e k o n a n i a c h d o t y c z ą c y c h f i l o z o f i c z n y c h , e t y c z n y c h , s o c j o l o g i c z n y c h i p s y c h o l o g i c z n y c h p o d s t a w e k o n o m i i ( U m i ń s k i 2 0 1 3 , s . 2 1 0 ) . D y s k u s j a n a t e m a t k o n s e k w e n c j i n i e r ó w n o ś c i s o c j e t e t n y c h – z w ł a s z c z a r o z w a r s t w i e n i a d o c h o d o w e g o – t o c z y s i ę n i e t y l k o n a g r u n t e n a u k s o c j e t e t n y c h , a l e t ą k ż e w d e b a c i e p u b l i c z n e j , w y w o ł u j ą c n i e r z a d k o w i e l e e m o c j i . N i e m n i e j j e d n a k i s t n i e j e p o w s z e c h n a z g o d a w ś r ó d b a d a c z y , ż e n a d m i e r n e n i e r ó w n o ś c i d o c h o d o w e n a r u s z a j ą z a s a d ę s p r a w i e d l i w o ś c i s o c j e t e t n e j o r a z w p ł y w a j ą n e g a t y w n i e n a r o z w ó j g o s p o d a r c z y ( P l i s z k a 2 0 0 4 , s . 3 5 4 ) . N i e r z a d k o w i e d e b a t a c h n a u k o w y c h i p o l i t y c z n y c h p o d k r e ś l a s i ę r ó w n i e ż , ż e p r e k r o c z e n i e p e w n e g o p r o g u r o z w a r s t w i e n i a d o c h o d ó w z a g r a ż a z a c h o w a n i u s p ó j n o ś c i s o c j e t e t n e j ( K o ł o d k o 2 0 1 4 , s . 3 2 ) . S p r a w ą k l u c z o w ą s t a j e s i ę z a t e m o k r e ś l e n i e , j a k i e n i e r ó w n o ś c i d o c h o d o w e n a l e ż y u z n a ć z a n a d m i e r n e , a j a k i e z a o p t y m a l n e . C e l e m p o l i t y k i s o c j e t e t n e j p o w i n n o b y ć b o w i e m n i w e l o w a n i e j e d y n i e n i e u z a s a d n i o n y c h , a n i e w s z e l k i c h n i e r ó w n o ś c i s o c j e t e t n y c h . W y d a j e s i ę , ż e p o m o c n e w t y m w z g l ę d z i e b ę d z i e r o z p a t r z e n i e k w e s t i i p r z y c z y n w y s t ę p u j ą c y c h n i e r ó w n o ś c i . O t ó ż ź r ó d ł e m r o z w a r s t w i e n i a d o c h o d o w e g o s o c j e t e t n s t w a s ą r ó ż n i c e w u w a r u n k o w a n i a c h ś r o d o w i s k o w y c h i g e n e t y c z n y c h o r a z r ó ż n i c e w p r e f e r e n c j a c h i a m b i c j a c h . Z m n i e j s z a n i e n i e r ó w n o ś c i w y n i k a j ą c y c h z t e j p i e r w s z e j g r u p y u w a r u n k o w a Ń j e s t b e z s p r z e c z n e – n i e w z b u d z a o n o w i e l u k o n t r o w e r s j i i w i ą z e s i ę z s z e r o k ą a k c e p t a c j ą s o c j e t e t n ą . Z k o l e i z m n i e j s z a n i e d y s p r o p o r c j i d o c h o d o w y c h w i ą z a n y c h z e z r ó ż n i c o w a n y m i d e c y z y j a m i j e d n o s t e k j e s t c o n a j m n i e j d y s k u s y j n e .

T o t e ż u p o d s t a w o d p o w i e d z i n a p y t a n i e , j a k i e n i e r ó w n o ś c i s o c j e t e t n e s ą s p r a w i e d l i w e , a j a k i e n i e , s t a ć p o w i n n o r o z r ó ż n i e n i e m i ę d z y d w o m a k a t e g o r i a m i – m o ż l i w o ś c i a m i i p r e f e r e n c j a m i . S p r a w i e d l i w e n i e r ó w n o ś c i t o t a k i e , z a k t ó r e o d p o w i e d z i a l n o ś ć p o n o s ą t y l k o i w y ł ą c z n i e j e d n o s t k i p o p r z e z s w o j e a u t o n o m i c z n e d e c y z j e – c z y t o e d u k a c y j n e , c z y z a w o d o w e , c z y t e z w i ą z a n e z e s t o p n i e m z a a n g a ż o w a n i a w p o p r a w ę s w o j e j s y t u a c j i b y t o w e j . N a t o m i a s t n i e s p r a w i e d l i w e n i e

równości społeczne to te niezależne od woli danej jednostki, na które nie ma ona żadnego wpływu, których nie kontroluje i nie może zmienić. Nie ulega wątpliwości, że czynniki takie, jak miejsce urodzenia, środowisko dorastania, sytuacja społeczno-ekonomiczna rodziców, najbliższe otoczenie, zdolności i predyspozycje, wpływają w dużej mierze na wielkość dochodu, jaki dana jednostka będzie osiągała w czasie swojego dorosłego życia, i powodują, że zasada równych szans zostaje zachwiana. Zatem aby dysproporcje w dochodzie były w pełni uzasadnione, należy przede wszystkim wyrównywać szanse. To natomiast, jak poszczególne osoby będą się zachowywać, zależy już wyłącznie od nich samych i tylko one ponoszą za swoje działania odpowiedzialność (Bartak 2014, s. 224). Za sprawiedliwą można bowiem uznać jedynie taką sytuację, gdy osobisty wysiłek decyduje o powodzeniu w życiu bardziej niż dziedziczone bogactwo czy korzystne środowisko rodzinne w dzieciństwie, które to środowisko wyposaża dziecko w odpowiedni kapitał kulturowy już na samym starcie i umożliwia mu dostęp do lepszej edukacji (Woźniak 2012, s. 27–28).

Przedmiotem analizy w niniejszym artykule są nierówności dochodowe wynikające z innych warunków życia w miastach i na wsiach. Nierówności tego typu należy niewątpliwie łagodzić poprzez stosowanie odpowiednio dobranych narzędzi. Dobrze zaprojektowana polityka społeczna powinna zatem ograniczać nierówności wynikające z faktu, że ludzie nie startują z tych samych pozycji w wyścigu o lepszą sytuację materialną, wyższy status społeczny i związane z nim ułatwienia. Najlepszym sposobem na redukcję nierówności dochodowych jest zapewnienie wszystkim grupom społecznym dostępu do nowoczesnej edukacji dostosowanej do wymagań gospodarki opartej na wiedzy. Konieczne jest również umożliwienie poszczególnym jednostkom dostępu do właściwej infrastruktury, korzystania ze zdobyczy postępu techniczno-technologicznego oraz całego spektrum osiągnięć cywilizacyjnych. Priorytetem polityki państwa powinno być zawsze dawanie równych szans, niwelowanie barier, stymulowanie innowacyjności oraz zapewnianie sprawiedliwej konkurencji. W nowoczesnej gospodarce polityka państwa nie może sprowadzać się zatem do działalności redystrybucyjnej, bo oczywiste jest, że to wzmocniałoby tylko postawy roszczeniowe, utrwałoby wyuczoną bezradność, ograniczało aktywność zawodową i samoopowiedzialność ludzi (Bartak 2014, s. 220). Właściwa polityka państwa w zakresie zmniejszania nierówności dochodowych nie spowalnia tempa procesów modernizacyjnych w gospodarce, a wręcz przeciwnie – prowadzi do ich przyspieszenia. Dysproporcje między terenami miejskimi i wiejskimi należy zatem łagodzić poprzez wspieranie dobrze zaprojektowanych inwestycji w kapitał ludzki oraz w poprawę infrastruktury.

Jak już wspomniano, poziom osiąganego dochodu warunkuje stopień zaspokojenia wielu różnych potrzeb człowieka, w tym potrzeb pozaekonomicznych. Wydaje się jednak, że w tym kontekście dodatkowego wyjaśnienia wymagają następujące trzy pojęcia: warunki życia, poziom życia i jakość życia.

Warunki życia obejmują nie tylko sytuację materialną człowieka, lecz również kwestie niematerialne, takie jak zdrowie, sytuacja zawodowa, poczucie integracji społecznej, formy spędzania wolnego czasu. Co istotne, warunki życia mogą być oceniane obiektywnie i subiektywnie. Poziom życia natomiast odnosi się wyłącz-

nie do stopnia zaspokojenia potrzeb konsumpcyjnych i opisywany jest za pomocą mierników obiektywnych (Słaby 2007, s. 104–105).

Podstawę opisu jakości życia stanowi z kolei subiektywna ocena stopnia zaspokojenia całego spektrum potrzeb – tych materialnych i tych niematerialnych. W rozważaniach dotyczących jakości życia uwagę zwraca się więc jednocześnie na takie aspekty życia, jak poczucie bezpieczeństwa (własnego, rodziny, własności, pracy), stan zdrowia, możliwości kształcenia się i rozwijania zainteresowań. A więc jakość życia zawiera te wszystkie elementy, które związane są z istnieniem, byciem kimś, odczuwaniem różnych stanów emocjonalnych (wypływających na przykład z faktu posiadania rodziny, przyjaciół, kolegów). W pojęciu jakości życia zawierają się też potrzeby związane z wartościami duchowymi, z zachowaniem godności, pragnieniem szacunku, poważania (Słaby 1990, s. 25). Oczywiście na zadowolenie danej osoby z poszczególnych obszarów własnego życia wpływ mają jej indywidualne aspiracje, hierarchia wartości, umiejętność radzenia sobie z problemami i przewycięzania trudności, odporność na stres, ogólne postrzeganie rzeczywistości itd. Oznacza to, że te same obiektywne warunki życia mogą być przez różne osoby zupełnie inaczej odbierane, stąd niewątpliwie inny będzie stopień zadowolenia tych osób z poszczególnych aspektów życia i inna też będzie ich subiektywna ocena dotycząca jakości życia.

## Rzeczowy, czasowy i przestrzenny zakres badania

Celem artykułu jest odpowiedź na pytanie, czy rozkład dochodu rozporządzalnego na osobę<sup>1</sup> na obszarach miejskich i wiejskich<sup>2</sup> jest taki sam, a jeśli jest inny, to jakie są między tymi rozkładami różnice. Aby udzielić odpowiedzi na tak zadane pytanie przeprowadzono badanie, w którym przeanalizowano oddzielnie piętnaście kolejnych lat od 1998 do 2012 r. Badanie to zrealizowano, korzystając z nieidentyfikowalnych danych jednostkowych z badania budżetów gospodarstw domowych<sup>3</sup> przez GUS<sup>4</sup>. W tabeli 1 zebrano informacje na temat struktury próby objętej analizą.

<sup>1</sup> Dochód rozporządzalny zdefiniowano za GUS jako sumę bieżących dochodów gospodarstwa domowego z poszczególnych źródeł pomniejszoną o zaliczki na podatek dochodowy od osób fizycznych płacone przez płatnika w imieniu podatnika, o podatki od dochodów z własności, podatki płacone przez osoby pracujące na własny rachunek oraz o składki na ubezpieczenia społeczne i zdrowotne. W skład dochodu rozporządzalnego wchodzi dochody pieniężne i niepieniężne, w tym spożycie naturalne (tj. towary i usługi konsumpcyjne pobrane na potrzeby gospodarstwa domowego z gospodarstwa indywidualnego w rolnictwie bądź prowadzonej działalności gospodarczej na własny rachunek) oraz towary i usługi otrzymane nieodpłatnie. Dochód rozporządzalny przeznaczony jest na wydatki oraz przyrost oszczędności (Budżety... 2015, s. 18, 19).

<sup>2</sup> Obszary wiejskie zdefiniowano za GUS jako wszystkie tereny położone poza granicami administracyjnymi miast (a więc obszary gmin wiejskich oraz części wiejskie – tj. leżące poza miastami – gmin miejsko-wiejskich).

<sup>3</sup> Trzeba jednak zaznaczyć, że za jednostkę statystyczną w niniejszym artykule przyjęto osobę, a nie gospodarstwo domowe.

<sup>4</sup> Bazę nieidentyfikowalnych danych jednostkowych z badania budżetów gospodarstw domowych za lata 1998–2012 udostępnił GUS na podstawie umowy nr 20/Z/DI-6-611/632/2013/RM między Głównym Urzędem Statystycznym i Uniwersytetem Szczecińskim.

Tabela 1. Struktura badanej próby

Lata	a* b c d	Miasta o liczbie mieszkańców					miasta razem	Wsie	Ogółem
		≥ 500 tys.	[200 tys., 500 tys.)	[100 tys., 200 tys.)	[20 tys., 100 tys.)	< 20 tys.			
1998	a	4439	3842	2490	6371	3898	21040	10716	31756
	b	14,0%	12,1%	7,8%	20,1%	12,3%	66,3%	33,7%	100,0%
	c	11654	10946	7301	19263	12348	61512	39186	100698
	d	11,5%	10,9%	7,3%	19,1%	12,3%	61,1%	38,9%	100,0%
1999	a	4388	3742	2486	6126	4022	20764	10664	31428
	b	14,0%	11,9%	7,9%	19,5%	12,8%	66,1%	33,9%	100,0%
	c	11338	10794	7371	18664	12530	60697	39031	99728
	d	11,4%	10,8%	7,4%	18,7%	12,6%	60,9%	39,1%	100,0%
2000	a	4197	4082	2986	7309	5027	23601	12562	36163
	b	11,6%	11,3%	8,3%	20,2%	13,9%	65,3%	34,7%	100,0%
	c	10865	11383	8574	22051	16063	68936	45349	114285
	d	9,5%	10,0%	7,4%	19,3%	14,1%	60,3%	39,7%	100,0%
2001	a	4289	3763	2601	6400	4005	21058	10789	31847
	b	13,5%	11,8%	8,2%	20,1%	12,5%	66,1%	33,9%	100,0%
	c	11081	10298	7383	18713	12760	60235	38943	99178
	d	11,2%	10,4%	7,4%	18,9%	12,8%	60,7%	39,3%	100,0%
2002	a	4412	3837	2806	6361	4125	21541	10801	32342
	b	13,5%	11,9%	8,7%	19,7%	12,8%	66,6%	33,4%	100,0%
	c	11223	10483	7788	18577	12862	60933	38883	99816
	d	11,2%	10,5%	7,8%	18,6%	12,9%	61,0%	39,0%	100,0%
2003	a	4570	3740	2951	6318	4232	21811	10641	32452
	b	14,1%	11,5%	9,1%	19,5%	13,0%	67,2%	32,8%	100,0%
	c	11600	10309	8203	18369	12938	61419	38080	99499
	d	11,6%	10,4%	8,2%	18,5%	13,0%	61,7%	38,3%	100,0%
2004	a	4600	3502	2945	6431	4228	21706	10508	32214
	b	14,3%	10,9%	9,1%	20,0%	13,1%	67,4%	32,6%	100,0%
	c	11695	9559	8373	18685	12852	61164	37801	98965
	d	11,8%	9,6%	8,5%	18,9%	13,0%	61,8%	38,2%	100,0%
2005	a	4486	3599	2703	6794	4001	21583	13184	34767
	b	12,9%	10,4%	7,8%	19,5%	11,5%	62,1%	37,9%	100,0%
	c	11180	9599	7533	19737	11955	60004	47120	107124
	d	10,4%	9,0%	7,0%	18,4%	11,2%	56,0%	44,0%	100,0%
2006	a	4559	3573	2854	6765	3945	21696	15812	37508
	b	12,2%	9,5%	7,6%	18,0%	10,5%	57,8%	42,2%	100,0%
	c	11187	9097	7949	19137	11675	59045	55266	114311
	d	9,8%	8,0%	7,0%	16,7%	10,2%	51,7%	48,3%	100,0%
2007	a	4574	3457	2916	6564	4097	21608	15758	37366
	b	12,2%	9,3%	7,8%	17,5%	11,0%	57,8%	42,2%	100,0%
	c	11262	8721	7820	18022	11957	57782	54210	111992
	d	10,0%	7,8%	7,0%	16,1%	10,7%	51,6%	48,4%	100,0%

Lata	a* b c d	Miasta o liczbie mieszkańców					miasta razem	Wsie	Ogółem
		≥ 500 tys.	[200 tys., 500 tys.)	[100 tys., 200 tys.)	[20 tys., 100 tys.)	< 20 tys.			
2008	a	4627	3424	2805	6716	4047	21619	15739	37358
	b	12,4%	9,2%	7,5%	18,0%	10,8%	57,9%	42,1%	100,0%
	c	10890	8670	7101	17930	11401	55992	53827	109819
	d	9,9%	7,9%	6,5%	16,3%	10,4%	51,0%	49,0%	100,0%
2009	a	4642	3482	2647	6592	4144	21507	15795	37302
	b	12,5%	9,3%	7,1%	17,7%	11,1%	57,7%	42,3%	100,0%
	c	10427	8704	6700	17493	11568	54892	53146	108038
	d	9,6%	8,1%	6,2%	16,2%	10,7%	50,8%	49,2%	100,0%
2010	a	4702	3432	2590	6343	4474	21541	15871	37412
	b	12,6%	9,2%	6,9%	16,9%	12,0%	57,6%	42,4%	100,0%
	c	10633	8471	6617	16917	12516	55154	52813	107967
	d	9,9%	7,8%	6,1%	15,7%	11,6%	51,1%	48,9%	100,0%
2011	a	4722	3367	2658	6435	4393	21575	15800	37375
	b	12,6%	9,0%	7,1%	17,2%	11,8%	57,7%	42,3%	100,0%
	c	10787	8284	6918	17004	12076	55069	52170	107239
	d	10,1%	7,7%	6,4%	15,9%	11,3%	51,4%	48,6%	100,0%
2012	a	4768	3455	2719	6447	4296	21685	15742	37427
	b	12,7%	9,2%	7,3%	17,2%	11,5%	57,9%	42,1%	100,0%
	c	10700	8396	6887	16733	11409	54125	51202	105327
	d	10,2%	8,0%	6,5%	15,9%	10,8%	51,4%	48,6%	100,0%

\* a – liczba przebadanych gospodarstw domowych, b – odsetek danych gospodarstw domowych w próbie, c – liczba przebadanych osób, d – odsetek danych osób w próbie

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

Populację generalną stanowią wszyscy mieszkańcy Polski. W tabeli 2 podano liczebność tej populacji w latach 1998–2012 oraz relację liczebności próby do liczebności zbiorowości generalnej.

Z danych znajdujących się w tabeli 2 wynika, że na przestrzeni lat 1998–2012 w Polsce nieznacznie zmniejszyła się liczba miast (o 329 miast). Z kolei liczba wsi wzrosła o 580. W 1998 r. w ramach badania budżetów gospodarstw domowych GUS przebadał co 380 osobę, w tym co 385 mieszkańca miast i co 372 mieszkańca wsi. Dla porównania – w 2012 r. ankietyzacji został poddany co 366 mieszkaniec Polski, w tym co 432 osoba z obszarów miejskich i co 296 osoba z obszarów wiejskich.

Badanie budżetów gospodarstw domowych przez GUS prowadzone jest metodą reprezentacyjną, która daje możliwość uogólnienia uzyskanych wyników na wszystkie gospodarstwa domowe w Polsce (Budżety... 2015, s. 14).

Dochód rozporządzalny przypadający na osobę wyznaczono w niniejszym artykule jako iloraz całkowitego dochodu gospodarstwa domowego i liczby jego członków. Założono zatem, że całkowity dochód gospodarstwa domowego jest dzielony po równo na wszystkich jego członków. Przyjęcie takiego upraszczającego założenia było możliwe, bowiem nie stało w sprzeczności z celami niniejszego

Tabela 2. Wielkość próby w stosunku do wielkości populacji generalnej

Lata	Liczebność zbiorowości generalnej (w tys.)*			Wielkość próby w stosunku do populacji generalnej (w %)		
	miasta	wsie	ogółem	miasta	wsie	ogółem
1998	23689	14594	38283	0,260	0,269	0,263
1999	23691	14579	38270	0,256	0,268	0,261
2000	23691	14565	38256	0,291	0,311	0,299
2001	23657	14594	38251	0,255	0,267	0,259
2002	23608	14624	38232	0,258	0,266	0,261
2003	23543	14652	38195	0,261	0,260	0,261
2004	23490	14690	38180	0,260	0,257	0,259
2005	23451	14710	38161	0,256	0,320**	0,281
2006	23401	14731	38132	0,252	0,375	0,300
2007	23351	14765	38116	0,247	0,367	0,294
2008	23305	14811	38116	0,240	0,363	0,288
2009	23294	14859	38153	0,236	0,358	0,283
2010	23448	15068	38517	0,235	0,350	0,280
2011	23405	15121	38526	0,235	0,345	0,278
2012	23360	15174	38534	0,232	0,337	0,273

\* dla poszczególnych lat podano stan na 30 czerwca, \*\*począwszy od 2005 r., dla uzyskania lepszej precyzji ocen dotyczących gospodarstw domowych rolników, zwiększono liczbę wiejskich terenowych punktów badań (Budżety... 2015, s. 16)

Źródło: jak w tabeli 1.

opracowania, jak również nie miało wpływu na ostateczne wnioski wynikające z przeprowadzonych analiz.

## Charakterystyka rozkładów za pomocą miar tendencji centralnej, zróżnicowania i asymetrii

Dla każdego gospodarstwa domowego ankietowanego przez GUS w ramach badania budżetów gospodarstw domowych wyznaczono miesięczny dochód rozporządzalny przypadający na osobę. Informacje zawarte w udostępnionej przez GUS bazie pozwoliły także na przyporządkowanie poszczególnych gospodarstw do jednej z dwóch rozłącznych i wyczerpujących grup – terenów miejskich i terenów wiejskich<sup>5</sup>. Następnie dla poszczególnych grup wyznaczono klasyczne miary tendencji centralnej, zróżnicowania i asymetrii. Obliczenia powtórzono piętnastokrotnie – dla każdego roku osobno od 1998 do 2012. Uzyskane wyniki zebrano w tabeli 3.

Zaobserwowano, że największe i najmniejsze wartości w poszczególnych szeregach statystycznych mają bardzo duży wpływ na obliczone w tabeli 3 klasyczne miary tendencji centralnej, zróżnicowania i asymetrii. Skłoniło to do refleksji,

<sup>5</sup> W artykule zostały one nazwane również miastami i wsiami.

Tabela 3. Wartości średniej arytmetycznej, odchylenia standardowego, klasycznego współczynnika zmienności oraz klasycznego współczynnika asymetrii

Lata	Średnia arytmetyczna (w zł)		Odchylenie standardowe (w zł)		Współczynnik zmienności (w %)		Współczynnik asymetrii	
	miasto	wieś	miasto	wieś	miasto	wieś	miasto	wieś
1998	578,75	408,58	395,35	332,08	68,3	81,3	5,52	6,53
1999	631,12	436,31	614,62	357,26	97,4	81,9	20,04	7,97
2000	685,58	477,71	489,46	668,54	71,4	139,9	5,18	42,15
2001	740,69	508,33	525,61	428,88	71,0	84,4	6,02	9,43
2002	769,90	522,96	584,95	713,73	76,0	136,5	11,55	50,18
2003	796,01	529,12	608,75	425,47	76,5	80,4	5,96	7,05
2004	824,45	544,09	625,51	561,74	75,9	103,2	4,80	-2,75
2005	849,53	581,44	664,55	583,08	78,2	100,3	8,99	-2,28
2006	935,34	653,14	721,39	558,14	77,1	85,5	7,35	6,11
2007	1045,80	742,94	883,79	744,00	84,5	100,1	14,16	22,05
2008	1197,71	841,17	939,93	1175,29	78,5	139,7	14,14	64,09
2009	1289,05	904,34	893,38	950,29	69,3	105,1	4,29	20,66
2010	1379,82	972,44	1381,76	951,94	100,1	97,9	58,87	13,77
2011	1428,72	998,15	1106,52	946,76	77,4	94,9	11,33	5,42
2012	1477,23	1065,17	1147,10	1101,08	77,7	103,4	11,42	12,14

Źródło: jak w tabeli 1.

że do porównania rozkładów dochodu rozporządzalnego na osobę w miastach i na wsiach należałoby wykorzystać szeregi pozbawione jednostek statystycznych o skrajnych wartościach badanej zmiennej. W celu podjęcia decyzji, jaki procent takich jednostek z poszczególnych prób odrzucić, wyznaczono:

- różnicę między wartością maksymalną  $x_{maks.}$  i minimalną  $x_{min.}$ ;
  - różnicę między wartością percentyla<sup>6</sup> dziewięćdziesiątego dziewiątego  $P_{99.100}$  i percentyla pierwszego  $P_{1.100}$ ;
  - różnicę między wartością decyla<sup>7</sup> dziewiątego  $D_{9.10}$  i decyla pierwszego  $D_{1.10}$ ;
  - różnicę między wartością kwartyla<sup>8</sup> trzeciego  $Q_{3.4}$  i kwartyla pierwszego  $Q_{1.4}$ .
- Wyniki przeprowadzonych obliczeń umieszczono w tabeli 4.

Na podstawie wyników zestawionych w tabeli 4 podjęto decyzję, że odrzucenie 1% jednostek statystycznych o największych wartościach badanej cechy oraz 1% jednostek o wartościach najmniejszych będzie wystarczające do wyeliminowania nadmiernego wpływu jednostek skrajnych na wartości obliczonych miar klasycznych. W tabeli 5 obliczono więc ponownie klasyczne miary tendencji centralnej, zróżnicowania i asymetrii, ale tym razem tylko dla 98% środkowych jednostek wyselekcjonowanych z poszczególnych prób.

Na podstawie liczb umieszczonych w tabeli 5 można stwierdzić, że w każdym roku z rozpatrywanych piętnastu lat średni dochód rozporządzalny na osobę w miastach był wyższy niż na wsiach. Odchylenie standardowe również miało większą wartość w przypadku miast niż wsi. Warto jednak dodać, że z wyjątkiem

<sup>6</sup> Percentyle dzielą zbiorowość statystyczną na sto równych części.

<sup>7</sup> Decyle dzielą zbiorowość statystyczną na dziesięć równych części.

<sup>8</sup> Kwartyle dzielą zbiorowość statystyczną na cztery równe części.



Tabela 4. Rozstęp obliczony dla wybranych obszarów zmienności

Lata	$x_{\max} - x_{\min}$ (w zł)		$P_{99,100} - P_{1,100}$ (w zł)		$D_{9,10} - D_{1,10}$ (w zł)		$Q_{3,4} - Q_{1,4}$ (w zł)	
	miasto	wieś	miasto	wieś	miasto	wieś	miasto	wieś
1998	11827,35	18232,58	1821,58	1562,16	706,23	531,58	344,76	265,35
1999	76320,30	15969,91	1932,00	1625,47	779,23	578,23	374,09	289,24
2000	22128,38	77468,79	2307,25	1801,36	884,67	635,48	423,98	323,37
2001	20617,27	24641,13	2395,23	2028,71	960,88	688,41	469,76	341,43
2002	35561,07	70970,47	2595,52	1870,57	1026,35	703,39	504,36	354,85
2003	23599,82	20022,08	2722,43	1959,46	1087,18	718,15	526,44	355,33
2004	21139,05	34896,58	2909,56	2116,40	1138,08	757,13	550,58	375,55
2005	42999,59	49218,81	3033,37	2280,39	1161,22	799,77	565,23	392,50
2006	38046,16	21875,17	3247,50	2553,12	1243,51	893,50	619,02	430,61
2007	50815,91	77449,48	3661,94	3042,89	1358,74	971,46	650,00	475,68
2008	80999,16	149215,95	3944,33	3276,52	1 530,29	1127,18	738,63	534,64
2009	32992,00	67633,74	4351,33	3711,67	1668,75	1216,42	797,44	587,93
2010	194790,15	51245,79	4618,33	3861,18	1779,30	1274,80	845,50	623,40
2011	72759,83	65426,65	4772,25	4218,95	1 828,40	1353,63	870,00	656,12
2012	72494,89	72263,10	5105,45	4485,25	1902,82	1419,95	917,00	694,00

Źródło: jak w tabeli 1.

2003 i 2004 r. różnicowanie dochodu wśród mieszkańców obszarów wiejskich było względnie większe niż wśród mieszkańców obszarów miejskich. Co ciekawe, po odrzuceniu 2% jednostek o skrajnych wartościach asymetria rozkładu dochodu w przypadku mieszkańców miast i wsi stała się zbliżona.

Tabela 5. Wartości średniej arytmetycznej, odchylenia standardowego, klasycznego współczynnika zmienności oraz klasycznego współczynnika asymetrii – obliczenia przeprowadzone dla zawężonego obszaru zmienności

Lata	Średnia arytmetyczna (w zł)		Odchylenie standardowe (w zł)		Współczynnik zmienności (w %)		Współczynnik asymetrii	
	miasto	wieś	miasto	wieś	miasto	wieś	miasto	wieś
1998	560,10	398,99	289,14	218,90	51,6	54,9	1,39	1,22
1999	610,52	424,15	315,97	235,82	51,8	55,6	1,34	1,18
2000	663,37	460,28	363,43	261,71	54,8	56,9	1,47	1,21
2001	717,27	493,48	392,03	284,38	54,7	57,6	1,38	1,35
2002	743,89	502,82	420,89	287,56	56,6	57,2	1,44	1,25
2003	768,21	514,42	445,61	294,61	58,0	57,3	1,44	1,29
2004	794,75	533,23	469,26	314,68	59,0	59,0	1,50	1,41
2005	818,96	565,54	480,37	337,66	58,7	59,7	1,52	1,47
2006	901,10	635,17	519,29	372,73	57,6	58,7	1,48	1,45
2007	1003,12	713,01	568,17	417,75	56,6	58,6	1,54	1,55
2008	1155,39	810,78	629,72	469,99	54,5	58,0	1,46	1,42
2009	1248,74	872,64	688,18	510,83	55,1	58,5	1,48	1,40
2010	1327,09	934,57	737,93	550,65	55,6	58,9	1,51	1,62
2011	1379,07	968,58	766,92	573,36	55,6	59,2	1,54	1,51
2012	1423,99	1024,65	792,31	613,49	55,6	59,9	1,52	1,56

Źródło: jak w tabeli 1.

Na podstawie oceny wartości wyznaczonych miar analizy struktury postawiono trzy hipotezy:

1. rozkład dochodu rozporządzalnego na mieszkańca na polskich wsiach jest inny niż w polskich miastach (tj. różnice w tych rozkładach są statystycznie istotne);
2. za to, że rozkłady te są inne, odpowiadają głównie różnice w przeciętnym poziomie dochodu w miastach i na wsiach (czyli różnice w zakresie dyspersji i skośności są dużo mniejsze niż różnice w zakresie tendencji centralnej);
3. na fakt, że rozkład dochodu rozporządzalnego na mieszkańca na terenach wiejskich nie jest taki sam jak na terenach miejskich, najmniejszy wpływ mają różnice w asymetrii (a zatem również i koncentracji) – wpływ różnic w zakresie asymetrii jest dużo mniejszy niż wpływ różnic w zakresie tendencji centralnej i zróżnicowania.

Do weryfikacji powyższych hipotez posłuży test Kołmogorowa-Smirnowa, który jest nieparametrycznym testem istotności i daje możliwość sprawdzenia, czy dwie populacje mają taki sam rozkład.

## Procedura realizacji stosowanego testu istotności

Niech rozpatrywana zmienna oznaczona zostanie przez  $X$ . Dystrybuanta  $F(X)$  w pełni określa rozkład zmiennej  $X$  w populacji. Stąd porównanie rozkładu zmiennej w dwóch populacjach można sprowadzić do porównania wartości dystrybuant w tych populacjach i jeżeli dwie populacje mają taki sam rozkład, to wartości ich dystrybuant powinny być we wszystkich punktach identyczne. Gdy więc populacja ludności miast oznaczona zostanie subskryptem  $M$ , a ludności wsi – subskryptem  $W$ , to aby udowodnić, że dwie populacje mają jednakowy rozkład, należy sprawdzić hipotezę zerową:

$$H_0: F_M(x_i) = F_W(x_i) \quad \text{dla każdej wartości zmiennej } X$$

wobec hipotezy alternatywnej:

$$H_1: F_M(x_i) \neq F_W(x_i) \quad \text{dla przynajmniej jednej wartości zmiennej } X,$$

gdzie  $i$  oznacza numer kolejnej obserwacji zmiennej  $X$ .

Wynika z tego, że jeśli dwie próby pochodzą z dwóch identycznych populacji, to wartości dystrybuant empirycznych (inaczej – doświadczalnych) obliczonych dla tych prób powinny być we wszystkich punktach zbliżone.

Niech próba pobrana z populacji ludności miast liczy  $n_M$  elementów, a z populacji ludności wsi –  $n_W$  elementów. Wówczas przez  $F_{n_M}(x_i)$  i  $F_{n_W}(x_i)$  oznaczone zostaną dystrybuanty empiryczne dotyczące odpowiednio pierwszej i drugiej próby. Przedmiotem analizy w rozpatrywanym teście są wielkości różnic pomiędzy wartościami tych dystrybuant. W celu określenia tych różnic wszystkie obserwacje występujące w próbach porządkuje się w kolejności niemalejącej. Następnie

dla każdej  $i$ -tej obserwacji oblicza się wartości obu dystrybuant empirycznych odpowiednio według wzorów:

$$F_{n_M}(x_i) = \frac{n_{Msk.}(x_i)}{n_M} \quad F_{n_W}(x_i) = \frac{n_{Wsk.}(x_i)}{n_W},$$

gdzie  $F_{n_M}(x_i)$  i  $F_{n_W}(x_i)$  są wartościami dystrybuant ustalonymi na podstawie próbek, a  $n_{Msk.}(x_i)$  i  $n_{Wsk.}(x_i)$  oznaczają liczebności skumulowane liczone odpowiednio dla pierwszej i drugiej próbki.

W następnym kroku przeprowadzania niniejszego testu szuka się największej wartości bezwzględnej różnicy między  $F_{n_M}(x_i)$  i  $F_{n_W}(x_i)$ . Znalazioną w ten sposób liczbę oznaczono przez  $D_{MW}$ . Tak więc wartość statystyki  $D_{MW}$  jest zdefiniowana jako (Kot i in. 2007, s. 267):

$$D_{MW} = \max_{x_i} |F_{n_M}(x_i) - F_{n_W}(x_i)|$$

Następnie w oparciu o statystykę  $D$  wyznacza się statystykę  $\lambda$  wyrażoną wzorem:

$$\lambda_{MW} = D_{MW} \sqrt{n_{MW}},$$

gdzie (Krysicki i in. 2003, s. 122):

$$n_{MW} = \frac{n_M n_W}{n_M + n_W}.$$

Z budowy statystyki  $\lambda$  wynika, że im większa będzie maksymalna różnica  $D$ , tym większą wartość będzie miała statystyka  $\lambda$  i tym większe będą podstawy do odrzucenia przypuszczenia o identyczności rozkładów w populacjach, z których wylosowano próby.

Przy założeniu prawdziwości hipotezy  $H_0$  statystyka  $\lambda$  – bez względu na postać poszczególnych dystrybuant – ma asymptotyczny rozkład  $\lambda$  Kołmogorowa (Józwiak, Podgórski 2012, s. 289). Z tablicy tego rozkładu dla przyjętego z góry poziomu istotności  $\alpha$  należy odczytać wartość krytyczną  $\lambda_\alpha$ , tak aby spełnione było równanie:

$$P\{\lambda \geq \lambda_\alpha\} = \alpha.$$

Następnie trzeba porównać obliczoną wartość  $\lambda$  z wartością krytyczną  $\lambda_\alpha$  i jeżeli zachodzi nierówność  $\lambda \geq \lambda_\alpha$ , to hipotezę  $H_0$  odrzuca się na rzecz hipotezy alternatywnej. Oznacza to wówczas, że dwie populacje, z których pochodzą próby, mają inny rozkład. Natomiast gdy spełniona jest nierówność  $\lambda < \lambda_\alpha$ , stwierdza się,

że brak jest podstaw do odrzucenia hipotezy zerowej o jednakowych rozkładach zmiennej  $X$  w obydwu populacjach.

## Zastosowanie testu Kołmogorowa-Smirnowa do badania zgodności rozkładów dochodu rozporządalnego na osobę mieszkańców miast i wsi

Opisana procedura testowania przeprowadzona będzie osobno dla każdego badanego roku. W pierwszym kroku obliczenia wykonane zostaną dla zmiennej  $X$ , czyli dla dochodu rozporządalnego na osobę. Na zaistniałe różnice w rozkładach będą miały wpływ wszystkie parametry tych rozkładów. Następnie te same obliczenia przeprowadzone zostaną dla zmiennej  $Y$  zdefiniowanej jako różnica między wartością zmiennej  $X$  oraz średnią arytmetyczną z  $X$  (tj.  $y_i = x_i - \bar{x}$ ). Wiadomo, że suma odchyłeń zmiennej  $X$  od jej średniej arytmetycznej zawsze wynosi 0. Oznacza to, że średnia dla rozkładu zmiennej  $Y$  będzie równa 0. Po przekształceniu zmiennej  $X$  na zmienną  $Y$  zróżnicowanie i asymetria rozkładu nie ulegną jednak zmianie, a zatem wartości odchylenia standardowego, klasycznego współczynnika zmienności oraz klasycznego współczynnika asymetrii będą jednakowe dla zmiennych  $X$  i  $Y$ .

W kolejnym kroku obliczenia przeprowadzone zostaną dla zmiennej  $Z$  oznaczającej zestandaryzowaną zmienną  $X$ , czyli  $z_i = \frac{x_i - \bar{x}}{S(x_i)}$ . Średnia dla rozpatrywanej

zmiennej po standaryzacji będzie wynosiła 0, natomiast odchylenie standardowe będzie równe 1. Co istotne, po przekształceniu zmiennej  $X$  w zmienną  $Z$  asymetria rozkładu pozostanie bez zmian, a zatem wartości klasycznego współczynnika asymetrii będą dla zmiennych  $X$  i  $Z$  jednakowe.

W tabeli 6 podano wartości statystyk  $D_{MW}$  i  $\lambda_{MW}$  uzyskane dla zmiennych  $X$ ,  $Y$  i  $Z$  (wszystkie obliczenia przeprowadzono dla zawężonego obszaru zmienności obejmującego 98% środkowych jednostek statystycznych).

Niech przyjęty z góry poziom istotności  $\alpha$  będzie równy 0,01. Wówczas odczytana z tablicy granicznego rozkładu  $\lambda$  Kołmogorowa wartość krytyczna dla założonego współczynnika  $\alpha = 0,01$  wynosi  $\lambda_\alpha = 1,63$ . Ponieważ dla zmiennej  $X$  otrzymano  $\lambda \geq \lambda_\alpha$ , wartość statystyki  $\lambda$  znalazła się w obszarze krytycznym i hipotezę  $H_0$  trzeba odrzucić (jest tak w przypadku wszystkich piętnastu lat). Stąd nie można twierdzić, że rozkład dochodu rozporządalnego na osobę w miastach i na wsiach jest taki sam – różnice między wartościami dystrybuant empirycznych w próbach były na tyle duże, że przypuszczenie o identyczności rozkładów w obu populacjach zostało odrzucone.

Hipotezę zerową trzeba było odrzucić także w przypadku porównywania rozkładów ujednoczonych pod względem tendencji centralnej, albowiem dla zmiennej  $Y$  też otrzymano  $\lambda \geq \lambda_\alpha$ . Również dla zmiennej  $Z$  wartość statystyki  $\lambda$  znalazła się w obszarze krytycznym i hipotezę  $H_0$  należało odrzucić (okazało się więc, że

Tabela 6. Wartości statystyk  $D_{MW}$  i  $\lambda_{MW}$  – obliczenia dla zmiennych  $X$ ,  $Y$  i  $Z$ 

Lata	Wyniki dla $x$		Wyniki dla $x - \bar{x}$		Wyniki dla $\frac{x - \bar{x}}{S(x)}$	
	$D_{MW}$	$\lambda_{MW}$	$D_{MW}$	$\lambda_{MW}$	$D_{MW}$	$\lambda_{MW}$
1998	0,2571	39,38	0,1106	16,93	0,0180	2,76
1999	0,2692	41,08	0,1152	17,58	0,0203	3,09
2000	0,2530	41,42	0,1266	20,73	0,0265	4,33
2001	0,2595	39,51	0,1274	19,40	0,0188	2,86
2002	0,2639	40,25	0,1505	22,95	0,0187	2,86
2003	0,2705	41,06	0,1629	24,72	0,0161	2,44
2004	0,2589	39,18	0,1606	24,30	0,0182	2,75
2005	0,2514	40,42	0,1561	25,11	0,0154	2,48
2006	0,2471	41,34	0,1475	24,68	0,0162	2,71
2007	0,2488	41,20	0,1398	23,15	0,0182	3,02
2008	0,2711	44,46	0,1296	21,26	0,0168	2,76
2009	0,2681	43,61	0,1311	21,33	0,0165	2,68
2010	0,2671	43,42	0,1326	21,55	0,0166	2,71
2011	0,2718	44,04	0,1206	19,54	0,0160	2,59
2012	0,2498	40,12	0,1143	18,36	0,0152	2,44

Źródło: jak w tabeli 1.

różnice między rozkładami ujednoczonymi pod względem tendencji centralnej i zróżnicowania były statystycznie istotne).

Skoro wiadomo już, że rozkłady w dwóch rozpatrywanych populacjach są inne, warto dodatkowo odpowiedzieć na pytanie, w jakim zakresie za istniejące rozbieżności odpowiadają różnice w tendencji centralnej, w jakim zakresie różnice w dyspersji, a w jakim różnice w asymetrii (a więc i koncentracji).

Dwa rozkłady są tym bardziej podobne, im mniejszą wartość ma  $\lambda$ . Wartość statystyki  $\lambda_{MW}$  obliczonej dla zmiennej  $Y$  jest o około połowę mniejsza od wartości tej statystyki obliczonej dla zmiennej  $X$ . Z kolei wartość statystyki  $\lambda_{MW}$  obliczonej dla zmiennej  $Z$  stanowi jedynie kilkanaście procent wartości tej statystyki obliczonej dla zmiennej  $Y$  oraz kilka procent wartości tej statystyki obliczonej dla zmiennej  $Z$ . Trzeba też wspomnieć o tym, że statystyka  $\lambda_{MW}$  obliczona dla zmiennej  $Z$  ma wartości bardzo bliskie wartości krytycznej  $\lambda_\alpha$ . Opisane prawidłowości dotyczą wszystkich badanych lat, stąd pozwalają na wyciągnięcie wniosku, że:

- za to, że rozkład dochodu rozporządkalnego na mieszkańca na polskich wsiach jest inny niż w polskich miastach odpowiada głównie różnica odnosząca się do tendencji centralnej;
- różnica w asymetrii ma dużo mniejszy – niż różnica w tendencji centralnej i dyspersji – wpływ na to, że rozkład dochodu rozporządkalnego na mieszkańca na terenach wiejskich jest inny niż na terenach miejskich;
- biorąc pod uwagę jedynie skośność, rozpatrywane rozkłady można uznać za bardzo podobne.

## Zalety proponowanej metody na tle metod stosowanych przez innych badaczy nierówności społecznych

Problem pomiaru nierówności dochodowych ma znaczące miejsce w literaturze ekonomicznej (Ulman, Wałęga 2006, s. 79). Autorki niniejszego artykułu zapoznały się z szeregiem prac poświęconych analizie dochodów różnych grup społecznych, a w szczególności z opracowaniami na temat metodyki badań różnic w dochodach mieszkańców miast i wsi. Literatura przedmiotu oferuje 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 na jakiejś jednej syntetycznej wielkości stanowiącej podstawę porównań. Przykładem takiego miernika nierówności jest współczynnik Theila, zgodnie z którym odmiennosc rozkładów opiera się wyłącznie na różnicy w wartościach średnich dla grup. Wzór na współczynnik Theila dotyczący nierówności między dochodami mieszkańców obszarów miejskich i mieszkańców obszarów wiejskich przedstawia się bowiem następująco (Allison 1978, s. 875):

$$T_{MW} = \frac{n_M \bar{x}_M}{n \bar{x}} \log_a \left( \frac{\bar{x}_M}{\bar{x}} \right) + \frac{n_W \bar{x}_W}{n \bar{x}} \log_a \left( \frac{\bar{x}_W}{\bar{x}} \right),$$

gdzie:

$T_{MW}$  – międzygrupowy współczynnik Theila,

$\bar{x}_M$  – wartość średnia w grupie gospodarstw domowych miejskich,

$\bar{x}_W$  – wartość średnia w grupie gospodarstw domowych wiejskich,

$\bar{x}$  – wartość średnia dla całej zbiorowości składającej się z dwóch grup,

$n_M$  – liczba przebadanych jednostek w grupie gospodarstw domowych z miast,

$n_W$  – liczba przebadanych jednostek w grupie gospodarstw domowych ze wsi,

$n$  – liczba wszystkich przebadanych jednostek statystycznych ( $n_M + n_W$ ),

$a$  – podstawa logarytmu, przy czym najczęściej wybieranymi logarytmami są logarytm dziesiętny (czyli logarytm o podstawie równej 10) oraz logarytm naturalny (czyli logarytm o podstawie równej stałej Eulera  $e$  wynoszącej w przybliżeniu 2,718).

Do analizy zróżnicowania międzygrupowego bardzo często wykorzystywana jest również wariancja międzygrupowa. Niestety wariancja międzygrupowa też, tak jak opisany powyżej międzygrupowy współczynnik Theila, opiera się jedynie na różnicy dotyczącej wartości średnich dla grup, pomijając różnice w dyspersji i asymetrii tych rozkładów. Sposób obliczenia wariancji opisującej dyspersję dochodu między grupą mieszkańców miast i grupą mieszkańców wsi prezentuje bowiem następująca formuła (Bellù, Liberati 2006a, s. 2, 10):

$$S_{MW}^2 = \frac{(\bar{x}_M - \bar{x})^2 n_M + (\bar{x}_W - \bar{x})^2 n_W}{n},$$

gdzie  $S_{MW}^2$  to wariancja międzygrupowa.

Współczynnik Theila bardzo często wykorzystywany jest przez badaczy zajmujących się porównywaniem różnych grup społecznych, ponieważ miara ta ma pewną przydatną własność. Otóż suma współczynnika międzygrupowego i średniego współczynnika wewnątrzgrupowego jest zawsze równa współczynnikowi ogólnemu obliczonemu dla całej badanej zbiorowości statystycznej (Hao, Naiman 2010, s. 50). Własność addytywnej dekomponowalności (Shorrocks 1980, s. 614) ma również drugi z przywołanych powyżej mierników, tj. wariancja (Domański i in. 2012, s. 132). Skoro więc wartość miernika ogółem jest sumą dwóch składników, to dzieląc każdy z tych składników przez wartość miernika ogółem, można obliczyć udziały składnika pierwszego i składnika drugiego w kształtowaniu się sumy (czyli udział nierówności międzygrupowych w nierównościach ogółem i udział nierówności wewnątrzgrupowych w nierównościach ogółem). Jeśli zatem celem badania jest ocena skali różnic między dochodami mieszkańców obszarów miejskich i wiejskich na tle takich różnic występujących wewnątrz grupy ludności miejskiej oraz wewnątrz grupy ludności wiejskiej, to współczynnik Theila i wariancja są narzędziami jak najbardziej przydatnymi (Bellù, Liberati 2006b, s. 7). Autorki niniejszego artykułu twierdzą jednak, że mierniki, które ujmują jedynie różnice między średnimi wartościami dla poszczególnych rozkładów, pomijając jednocześnie różnice w kształtach tych rozkładów (tj. w zróżnicowaniu i asymetrii), nie są narzędziami odpowiednimi do spełnienia postawionego w artykule celu kompleksowego porównania rozkładów i wskazania przyczyn istniejących między tymi rozkładami różnic. Zaproponowana w pracy metodyka badawcza jest innowacyjna i ma bardzo duży walor aplikacyjny.

Zaletą proponowanej metody jest również to, że nie wymaga ona przekształcenia szeregu szczegółowych w szeregi rozdzielcze przedziałowe. Jest to bardzo istotny walor, ponieważ wspomniane przekształcenie spowodowałoby utratę dużej ilości informacji o danych pierwotnych i byłoby daleko idącym kompromisem. Autorki artykułu chciały wykorzystać wszystkie informacje dostępne w szeregach wyliczających i nie stosować uproszczeń, które wpłynęłyby na wyniki końcowe przeprowadzonej analizy.

## Podsumowanie

Cel zintegrowanego rozwoju kraju wymaga łagodzenia różnic w poziomie życia w poszczególnych jego regionach. Wszystkie wymiary spójności – tj. kohezję społeczną, gospodarczą i terytorialną – należy rozpatrywać łącznie, bowiem między wymiarami tymi następują wzajemne interakcje. Nadmierne nierówności w poziomie życia w regionach wpływają na osłabienie spójności społeczno-gospodarczej całej tkanki kraju i są barierą harmonijnego rozwoju.

Idea spójności w wymiarze społecznym, gospodarczym i terytorialnym wpisana jest w strategię zrównoważonego rozwoju Unii Europejskiej i jest jednym z priorytetów polityki krajowej (Wołoszyn, Głowacka-Wołoszyn 2015, s. 397). Art. 2 ustawy z dnia 6 grudnia 2006 r. o zasadach prowadzenia polityki rozwoju kraju określa tę politykę jako „zespół wzajemnie powiązanych działań podejmo-

wanych i realizowanych w celu zapewnienia trwałego i zrównoważonego rozwoju kraju, spójności społeczno-gospodarczej, regionalnej i przestrzennej, podnoszenia konkurencyjności gospodarki oraz tworzenia nowych miejsc pracy w skali krajowej, regionalnej lub lokalnej”.

Nie ulega wątpliwości, że nadmierne różnice w poziomie życia ludności miast i wsi są przejawem braku spójności i stanowią zagrożenie dla zrównoważonego rozwoju kraju. Państwo musi więc stworzyć lepsze warunki instytucjonalne dla rozwoju obszarów wiejskich (Polski 2015, s. 78). Winno jednak stosować takie rozwiązania, które równocześnie nie godzą w efektywność gospodarowania (Bartak 2014, s. 231).

Obowiązującym dokumentem określającym cele i sposób działania podmiotów publicznych w odniesieniu do polskiej przestrzeni dla osiągnięcia strategicznych celów rozwoju Polski jest „Krajowa Strategia Rozwoju Regionalnego 2010–2020: regiony, miasta, obszary wiejskie” (KSRR). Dokument ten wyznacza cele polityki rozwoju regionalnego – w tym wobec obszarów wiejskich i miejskich – oraz określa instrumenty ich realizacji. Przyjęta polityka zakłada specjalne działania dla terytoriów wymagających szczególnego wsparcia w zakresie procesów rozwojowych, które to terytoria uznano za obszary problemowe. Do obszarów problemowych zaliczono tereny wiejskie, na których ma być ułatwiony dostęp do usług publicznych i podnoszona ich jakość. KSRR określa również zasady i mechanizmy współpracy oraz koordynacji działań podejmowanych na poziomie krajowym z działaniami podejmowanymi przez samorząd terytorialny na poziomie województw oraz przez pozostałych uczestników polityki regionalnej. Szczegółowe propozycje rozwiązań dotyczą przede wszystkim „(...) zwiększenia spójności terytorialnej – zarówno w skali krajowej, jak również regionalnej – przez budowanie powiązań funkcjonalnych między miastami wojewódzkimi a ich otoczeniem regionalnym, a zwłaszcza między miastami i obszarami wiejskimi, a także przez aktywne przeciwdziałanie marginalizacji najsłabiej rozwijających się obszarów (...)” (KSRR, s. 1378–1379).

Celem artykułu była odpowiedź na pytanie, czy w Polsce rozkład dochodu rozporządzalnego na mieszkańca na obszarach wiejskich jest taki sam jak na obszarach miejskich, a jeśli jest inny, to jakie są w tym zakresie różnice. Aby osiągnąć postawiony cel, sprawdzono, czy rozbieżności w wartościach dystrybuant obliczonych na podstawie wyników z prób są statystycznie istotne. Obliczenia zrealizowano po wyeliminowaniu z poszczególnych prób 2% jednostek o skrajnych wartościach badanej cechy. Weryfikacji odpowiedniej hipotezy statystycznej dokonano za pomocą testu Kołmogorowa-Smirnowa. Testowanie przeprowadzono dla piętnastu kolejnych lat od 1998 do 2012 r. łącznie. Badanie wykonano również:

- dla rozkładów ujednoczonych pod względem tendencji centralnej, tak aby określić łączny wpływ zróżnicowania i asymetrii na różnice w rozkładach;
- dla rozkładów ujednoczonych pod względem tendencji centralnej i zróżnicowania, tak aby określić wpływ asymetrii (a zatem i koncentracji) na różnice w rozkładach.



W każdym z 45 przypadków procedury testowania przypuszczenie o identyczności rozkładów dochodu rozporządzalnego na osobę w miastach i na wsiach trzeba było odrzucić, gdyż wartość statystyki empirycznej  $\lambda$  znalazła się w prawostronnym obszarze krytycznym określonym równaniem:  $P\{\lambda \geq 1,63\} = 0,01$ .

Reasumując, należy stwierdzić, że w Polsce średni dochód rozporządzalny na osobę jest znacznie większy w miastach niż na wsiach i głównie przez ten fakt rozkład tego dochodu w obydwu populacjach nie jest identyczny. Natomiast zróżnicowanie można uznać za podobne w obydwu populacjach. Z kolei asymetria w przypadku tych rozkładów jest na tyle podobna, że można ją nazwać prawie identyczną.

Zastosowany w artykule podział na obszary miejskie i obszary wiejskie jest podziałem z formalnego punktu widzenia poprawnym, lecz może budzić pewne zastrzeżenia natury merytorycznej. Kategoria 'miasto' oraz kategoria 'wieś' są bowiem wewnątrznie zróżnicowane i zróżnicowanie to dotyczy nie tylko liczby ludności, ale również poziomu rozwoju ekonomicznego oraz zakresu realizowanych wobec mieszkańców funkcji. Należy więc mieć świadomość faktu, że wyciągnięte w artykule wnioski oparto na porównaniu dwóch agregatów wewnątrznie bardzo niejednorodnych i stąd wyniki niniejszych badań tłumaczą przestrzenny wymiar nierówności społecznych w Polsce tylko w ograniczonym zakresie.

## Literatura

- Allison P.D. 1978. Measures of inequality. *American Sociological Review*, 43. USA.
- Bal I., 2012. Marginalizacja i wykluczenie społeczne jako bariera rozwoju regionalnego. *Nierówności Społeczne a Wzrost Gospodarczy*, 28. Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów.
- Bartak J. 2014. Uwarunkowania redukcji nierówności dochodowych w Polsce. *Nierówności Społeczne a Wzrost Gospodarczy*, 37. Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów.
- Bellù L.G., Liberati P. 2006a. Policy impacts on inequality: decomposition of income inequality by subgroups. FAO – Food and Agriculture Organization of the United Nations, Włochy.
- Bellù L.G., Liberati P. 2006b. Policy impacts on inequality: inequality and axioms for its measurement. FAO – Food and Agriculture Organization of the United Nations, Włochy.
- Budżety gospodarstw domowych w 2014 r. 2015. Główny Urząd Statystyczny, Warszawa.
- Domański H., Karpiński Z., Pokropek A., Przybysz D., Sawiński Z., Słomczyński K.M., Trzciniński R. 2012. Metodologia badań nad stratyfikacją społeczną. Wydawnictwo Naukowe SCHOLAR, Warszawa.
- Hao L., Naiman D.Q. 2010. *Assessing inequality*. SAGE Publications.
- Jóźwiak J., Podgórski J. 2012. *Statystyka od podstaw*. Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Kołodko G. 2014. Społeczne i przestrzenne aspekty zróżnicowania dochodów we współczesnym świecie. *Nierówności Społeczne a Wzrost Gospodarczy*, 39. Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów.
- Kot S.M., Jakubowski J., Sokołowski A. 2007. *Statystyka*. Wydawnictwo DIFIN, Warszawa.
- Krajowa Strategia Rozwoju Regionalnego 2010–2020: regiony, miasta, obszary wiejskie (M.P. z 2011 nr 36, poz. 423).
- Krysicki W., Bartos J., Dyczka W., Królikowska K., Wasilewski M. 2003. *Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. Cz. II. Statystyka matematyczna*. Wydawnictwo Naukowe PWN, Warszawa.
- Leszczyńska M. 2014. Ocena społecznego zrównoważenia rozwoju w Polsce według kryterium dynamiki dochodów gospodarstw domowych. *Nierówności społeczne a wzrost gospodarczy*, 37. Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów.

- Pliszka T. 2004. Skutki nierówności społecznych. *Nierówności Społeczne a Wzrost Gospodarczy*, 5. Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów.
- Polski J. 2015. Spójność terytorialna jako podstawa polityki regionalnej. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 392.
- Shorrocks A.F. 1980. The class of additively decomposable inequality measures. *Econometrica* 48, 3.
- Słaby T. 1990. Poziom życia, jakość życia. *Wiadomości Statystyczne*, 6.
- Słaby T. 2007. Poziom i jakość życia. [W:] T. Panek (red.), *Statystyka społeczna*. Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Ulman P., Wałęga A. 2006. Nierówności dochodowe w Polsce i ich dekompozycja. *Zeszyty Naukowe Polskiego Towarzystwa Ekonomicznego*, 4.
- Umiński P. 2013. Nierówności dochodowe w koncepcji Johna K. Galbraitha – wskazanie źródeł i sformułowanie hipotez badawczych. *Nierówności Społeczne a Wzrost Gospodarczy*, 30. Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów.
- Ustawa z dnia 6 grudnia 2006 r. o zasadach prowadzenia polityki rozwoju (tekst jednolity: Dz.U. 2016, poz. 383).
- Wołoszyn A., Głowacka-Wołoszyn R. 2015. Nierówności dochodowe gospodarstw domowych w Polsce w kontekście zrównoważonego rozwoju. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 395.
- Woźniak W. 2012. Nierówności społeczne w polskim dyskursie politycznym. Wydawnictwo Naukowe Scholar, Warszawa.
- Wójcik-Żołądek M. 2013. Nierówności społeczne w Polsce. *INFOS Zagadnienia Społeczno-Gospodarcze*, 20(157). Biuro Analiz Sejmowych.

## Application of Kolmogorov-Smirnov test to examining equality of inhabitants' income distributions in rural and urban areas in Poland

**Abstract:** The aim of the article is to answer the question whether the distribution of disposable income *per capita* for inhabitants of Polish cities and towns is the same as for inhabitants of Polish villages, and if not, what are the differences. To verify appropriate statistical hypotheses Kolmogorov-Smirnov test was used. The calculations were made separately for fifteen consecutive years. Results of the study indicate that the average disposable income *per capita* is much higher in urban areas than in rural areas and these differences in central tendency are the main reason that distributions considered cannot be regarded as identical. However, it can be assumed that dispersion of income in Polish villages and Polish cities/towns is similar and the scewness of these distributions is nearly identical.

**Key words:** disposable income *per capita*, distribution, Kolmogorov-Smirnov test