

Uniwersytet
Ekonomiczny
w Krakowie

Zeszyty Naukowe

Cracow Review
of Economics
and Management

923

Metody analizy danych

Kraków 2013

Rada Naukowa

Andrzej Antoszewski (Polska), Slavko Arsovski (Serbia), Josef Arlt (Czechy), Daniel Baier (Niemcy), Hans-Hermann Bock (Niemcy), Ryszard Borowiecki (Polska), Giovanni Lagioia (Włochy), Tadeusz Markowski (Polska), Martin Mizla (Słowacja), David Ost (USA), Józef Pociecha (Polska)

Komitet Redakcyjny

Paweł Lula, Barbara Pawełek (sekretarz), Adam Sagan, Tadeusz Sikora, Edward Smaga, Andrzej Sokołowski (przewodniczący)

Redaktor statystyczny

Andrzej Sokołowski

Redaktorzy Wydawnictwa

Monika Rusin, Seth Stevens (streszczenia w j. angielskim)

Projekt okładki i układ graficzny tekstu

Marcin Sokołowski

Streszczenia artykułów są dostępne w międzynarodowej bazie danych The Central European Journal of Social Sciences and Humanities <http://cejsh.icm.edu.pl> oraz w Central and Eastern European Online Library www.ceeol.com, a także w adnotowanej bibliografii zagadnień ekonomicznych i pokrewnych BazEkon http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php

© Copyright by Uniwersytet Ekonomiczny w Krakowie, Kraków 2013

ISSN 1898-6447

Wersja pierwotna: publikacja drukowana
Publikacja jest dostępna w bazie CEEOL (www.ceeol.com)
oraz w czytelni on-line ibuk.pl (www.ibuk.pl)

Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie
31-510 Kraków, ul. Rakowicka 27, tel. 12 293 57 42, e-mail: wydaw@uek.krakow.pl
www.zeszyty-naukowe.uek.krakow.pl

Zakład Poligraficzny Uniwersytetu Ekonomicznego w Krakowie
31-510 Kraków, ul. Rakowicka 27

Objętość 6,2 ark. wyd.
Zam. 104/2014

Spis treści

Zofia Mielecka-Kubień, Katarzyna Warzecha	
Sytuacja demograficzno-społeczna wybranych miast województwa śląskiego w latach 2002 i 2009	5
Marcin Salamaga	
Propozycja modyfikacji wskaźnika ujawnionej przewagi komparatywnej z wykorzystaniem funkcji wymiernych	23
Jan Tatar	
Modele wskaźnikowe rynku kapitałowego wykorzystujące funkcję regresji wektorów losowych	37
Katarzyna Budny	
Wybrane własności kurtozy wektora losowego	47
Marta Targaszewska	
Metody pomiaru jakości kształcenia na uczelniach wyższych	59
Jacek Stelmach	
O porównaniu dwóch populacji	71
Kamil Fijorek	
Aproksymacja modelu regresji logistycznej Firtha za pomocą ważenia obserwacji	87
Bogumiła Krzeszowska-Zakrzewska	
Zastosowanie programowania zero-jedynkowego w harmonogramowaniu czynności projektu	99
Lista recenzentów Zeszytów Naukowych Uniwersytetu Ekonomicznego w Krakowie opublikowanych w 2013 roku	119

Zofia Mielecka-Kubiś

Katarzyna Warzecha

Katedra Ekonometrii

Uniwersytet Ekonomiczny w Katowicach

Sytuacja demograficzno-społeczna wybranych miast województwa śląskiego w latach 2002 i 2009

Streszczenie

Celem badań jest określenie podobieństw miast województwa śląskiego liczących więcej niż 100 tys. mieszkańców (Bielsko-Biała, Bytom, Chorzów, Częstochowa, Dąbrowa Górnicza, Gliwice, Katowice, Ruda Śląska, Rybnik, Sosnowiec, Tychy i Zabrze), położonych w niezbyt dużej odległości geograficznej, w odstępie kilku lat. Podobieństwo badano z uwzględnieniem wybranych mierników sytuacji społeczno-gospodarczej miast i współczynników demograficznych w latach 2002 i 2009. Zastosowano metodę J. Czekańskiego, a w celu stwierdzenia, które grupy miast znajdują się w lepszej lub gorszej sytuacji pod jakimś względem, porównano średnie wybranych mierników ogółem i dla wyróżnionych grup miast.

Badania przeprowadzono w następujących przekrojach: I – sytuacja demograficzno-społeczna, II – poziom zdrowotności i niektóre jego uwarunkowania, III – poziom zamożności, IV – poziom kultury i edukacji.

Najbardziej podobnymi do siebie miastami, a jednocześnie będącymi w najlepszej sytuacji społeczno-gospodarczej i demograficznej, były zarówno w 2002 r., jak i w 2009 r., mimo dzielącej je sporej odległości geograficznej, Bielsko-Biała i Częstochowa, przy czym miasto Bielsko-Biała cechowało się wysokim poziomem zamożności, nie wykazując pod tym względem podobieństwa do innych miast. Najbardziej podobnymi do siebie, a zarazem będącymi w niekorzystnej sytuacji miastami okazały się: Sosnowiec, Zabrze i Gliwice, leżące na terenie byłego Górnośląskiego Okręgu Przemysłowego.

Słowa kluczowe: metoda Czekanowskiego, sytuacja demograficzno-społeczna, miasta województwa śląskiego, metody taksonomiczne.

1. Wprowadzenie

W województwie śląskim jest obecnie 12 miast o liczbie ludności przekraczającej 100 tys. (Bielsko-Biała, Bytom, Chorzów, Częstochowa, Dąbrowa Górnicza, Gliwice, Katowice, Ruda Śląska, Rybnik, Sosnowiec, Tychy i Zabrze). Miasta te odgrywają ważną rolę w rozwoju społeczno-ekonomicznym regionu, oddziałują one także w różny sposób na otaczające je mniejsze miasta i wsie. Trzy z wymienionych miast (Częstochowa, Gliwice i Sosnowiec) liczyły w badanym okresie ponad 200 tys. mieszkańców, a stolica regionu – Katowice, ponad 300 tys. Celem prezentowanych badań jest określenie podobieństw miast liczących więcej niż 100 tys. mieszkańców, położonych w obrębie tego samego województwa, a więc w niezbyt dużej odległości geograficznej, w odstępie kilku lat. Podobieństwa badano, biorąc pod uwagę mierniki opisujące sytuację społeczno-gospodarczą miast i współczynniki demograficzne w latach 2002 i 2009.

Zastosowano należącą do grupy metod analizy wielowymiarowej metodę J. Czekanowskiego, której graficzną prezentacją jest diagram Czekanowskiego¹. Pozwala on dokonać podziału obiektów (w tym przypadku miast) na grupy charakteryzujące się dużym stopniem wzajemnego podobieństwa, a otrzymane wyniki nie mają charakteru wartościującego, tzn. nie wskazują, które z grup obiektów są lepsze czy gorsze pod jakimś względem, informują jedynie o stopniu podobieństwa obiektów. Aby stwierdzić, które grupy miast znajdują się w lepszej lub gorszej sytuacji pod jakimś względem, porównano średnie ogólne i wyróżnionych grup miast dla branych pod uwagę mierników.

Odległość między badanymi obiektami w przestrzeni wielowymiarowej określano jako odległość euklidesową, a w przypadku grup częściowo pokrywających się wyboru odpowiedniego sposobu podziału miast dokonano, kierując się kryterium najkrótszej, dla możliwych grup przynależności danego miasta, przeciętnej odległości euklidesowej niejednoznacznie określonego miasta od innych miast.

Sytuację demograficzną, społeczną i ekonomiczną ludności miast województwa śląskiego (o liczbie mieszkańców powyżej 100 tys.) scharakteryzowano za pomocą zmiennych diagnostycznych opisujących różne jej aspekty² (tabela 1) i spełniających kryteria formalne i merytoryczne dotyczące zastosowania metody

¹ Diagram Czekanowskiego jest najstarszą metodą taksonomiczną, zastosowaną po raz pierwszy w 1909 r. przez polskiego antropologa J. Czekanowskiego. Więcej informacji na temat tej metody czytelnik może znaleźć w pracach: [Pluta 1986, Pocięcha i in. 1988, Wolny 2009, Czekanowski 1909 i 1913, Heffner i Gibas 2007]. Do obliczeń wykorzystano program komputerowy MaCzek.

² Wybrane zmienne spełniają wymagane kryteria formalne i merytoryczne: są ważne z punktu widzenia prowadzonych badań, słabo skorelowane między sobą oraz cechują się odpowiednio wysoką zmiennością (współczynnik zmienności powyżej 10%).

Czekanowskiego; zmienne te zostały doprowadzone do postaci umożliwiającej ich porównywalność:

- destymulanty³ (D) zamieniono na stymulanty (S) według kryterium:

$$D \rightarrow S = x_{\max} - x_i, \quad (1)$$

przy czym x_{\max} oznacza najwyższą zaobserwowaną wartość danej zmiennej;

- nominaty zamieniono na stymulanty (por. [Lipieta i in. 2000, s. 81]) według wzoru:

$$s_{ij} = -|N_{nom} - x_{ij}|, \quad (2)$$

gdzie N_{nom} – oznacza pożądaną wartość nominanty;

- wszystkie zmienne diagnostyczne zestandaryzowano według formuły:

$$x_{ij} = \frac{x_j - \bar{x}_j}{s_j}, \quad (3)$$

gdzie:

x_{ij} – wartość j -tej zmiennej w i -tym obiekcie (mieście),

\bar{x}_j – średnia arytmetyczna j -tej zmiennej,

s_j – odchylenie standardowe j -tej zmiennej ($i = 1, 2, \dots, n; j = 1, 2, \dots, m$).

Dane statystyczne pochodzą z Banku Danych Regionalnych GUS w Warszawie.

2. Wyniki badania empirycznego

W tabeli 1 przedstawiono zmienne diagnostyczne⁴ charakteryzujące różne dziedziny społecznej, demograficznej i ekonomicznej sytuacji ludności badanych miast.

Tabela 1. Zbiór cech diagnostycznych

Nazwa zmiennej	Symbol	
	2002	2009
Pracujący ogółem na 1000 ludności (S)	X_{12}	X_{19}
Liczba bezrobotnych na 1000 pracujących (D)	X_{22}	X_{29}
Przeciętne miesięczne wynagrodzenie brutto (S)	X_{32}	X_{39}

³ Zmienna (X) jest: stymulantą – im wyższa jest wartość tej zmiennej, tym lepsza jest sytuacja danego obiektu, nominantą – zmienna ma określoną najkorzystniejszą wartość zwaną wartością nominalną, lub destymulantą – im wyższa jest wartość tej zmiennej, tym gorsza jest sytuacja danego obiektu.

⁴ Ostatnia cyfra subskrypty zmiennej oznacza rok (2 – 2002, 9 – 2009). Szczegółowy opis badań można znaleźć w pracy [Mielecka-Kubień i Warzecha 2010].

cd. tabeli 1

Nazwa zmiennej	Symbol	
	2002	2009
Dochody budżetu miasta na jednego mieszkańca w zł (<i>S</i>)	X_{42}	X_{49}
Wydatki budżetu miasta na jednego mieszkańca w zł (<i>S</i>)	X_{52}	X_{59}
Podmioty zarejestrowane w rejestrze Regon na 10 tys. ludności (<i>S</i>)	X_{62}	X_{69}
Nakłady inwestycyjne na jednego mieszkańca w zł (<i>S</i>)	X_{72}	X_{79}
Wartość brutto środków trwałych na jednego mieszkańca w zł (<i>S</i>)	X_{82}	X_{89}
Ścieki oczyszczane w % wymagających oczyszczenia (<i>S</i>)	X_{92}	X_{99}
Emisja zanieczyszczeń gazowych w t/km ² (<i>D</i>)	X_{102}	X_{109}
Emisja zanieczyszczeń pyłowych w t/km ² (<i>D</i>)	X_{112}	X_{119}
Zużycie wody w gospodarstwach domowych na jednego mieszkańca w m ³ (<i>S</i>)	X_{122}	X_{129}
Zużycie gazu w gospodarstwach domowych na jednego mieszkańca w m ³ (<i>S</i>)	X_{132}	X_{139}
Mieszkania oddane do użytku na 1000 ludności (<i>S</i>)	X_{142}	X_{149}
Liczba ludności na jednego lekarza (<i>D</i>)	X_{152}	X_{159}
Liczba ludności na jednego dentystę (<i>D</i>)	X_{162}	X_{169}
Liczba ludności na jedną pielęgniarkę (<i>D</i>)	X_{172}	X_{179}
Liczba ludności na jedną aptekę (<i>S</i>)	X_{182}	X_{189}
Łóżka w szpitalach ogólnych na 10 tys. ludności (<i>S</i>)	X_{192}	X_{199}
Liczba zgonów niemowląt na 1000 urodzeń żywych (<i>D</i>)	X_{202}	X_{209}
Liczba ludności na jedną placówkę biblioteczną (<i>S</i>)	X_{212}	X_{219}
Księgozbiór na 1000 ludności w woluminach (<i>S</i>)	X_{222}	X_{229}
Liczba czytelników na 1000 ludności (<i>S</i>)	X_{232}	X_{239}
Liczba ludności na jedno miejsce w kinach stałych (<i>S</i>)	X_{242}	X_{249}
Liczba uczniów szkół podstawowych przypadających na jeden komputer z dostępem do internetu (<i>D</i>)	X_{252}	X_{259}
Liczba uczniów gimnazjum przypadających na jeden komputer z dostępem do internetu (<i>D</i>)	X_{262}	X_{269}
Liczba uczniów liceów ogólnokształcących na 1000 ludności (<i>S</i>)	X_{272}	X_{279}
Liczba uczniów ponadgimnazjalnych zasadniczych szkół zawodowych na 1000 ludności (<i>S</i>)	X_{282}	X_{289}
Liczba samochodów osobowych na 1000 ludności (<i>S</i>)	X_{292}	X_{299}
Drogi publiczne powiatowe o twardej nawierzchni (na 100 km ² powierzchni) w km (<i>S</i>)	X_{302}	X_{309}
Gęstość zaludnienia (<i>D</i>)	X_{312}	X_{319}
Udział ludności w wieku przedprodukcyjnym (<i>S</i>)	X_{322}	X_{329}

cd. tabeli 1

Nazwa zmiennej	Symbol	
	2002	2009
Współczynnik feminizacji (N)	X_{332}	X_{339}
Współczynnik przyrostu naturalnego (S)	X_{342}	X_{349}
Standaryzowany współczynnik małżeństw (S)	X_{352}	X_{359}
Standaryzowany współczynnik rozwodów (D)	X_{362}	X_{369}
Standaryzowany współczynnik płodności kobiet (S)	X_{372}	X_{379}
Standaryzowany współczynnik zgonów mężczyzn (D)	X_{382}	X_{389}
Standaryzowany współczynnik zgonów kobiet (D)	X_{392}	X_{399}
Współczynnik salda migracji (S)	X_{402}	X_{409}

Źródło: opracowanie własne.

Badania przeprowadzono dla całego zbioru współczynników (tabela 1) oraz następujących ich podzbiorów, charakteryzujących różne dziedziny społecznej, demograficznej i ekonomicznej sytuacji ludności wyróżnionych miast:

I – wszystkie zmienne,

II – zmienne charakteryzujące poziom zdrowotności i niektóre jego uwarunkowania w badanych miastach: X_9 – X_{11} , X_{15} – X_{20} , X_{31} , X_{38} – X_{39} ,

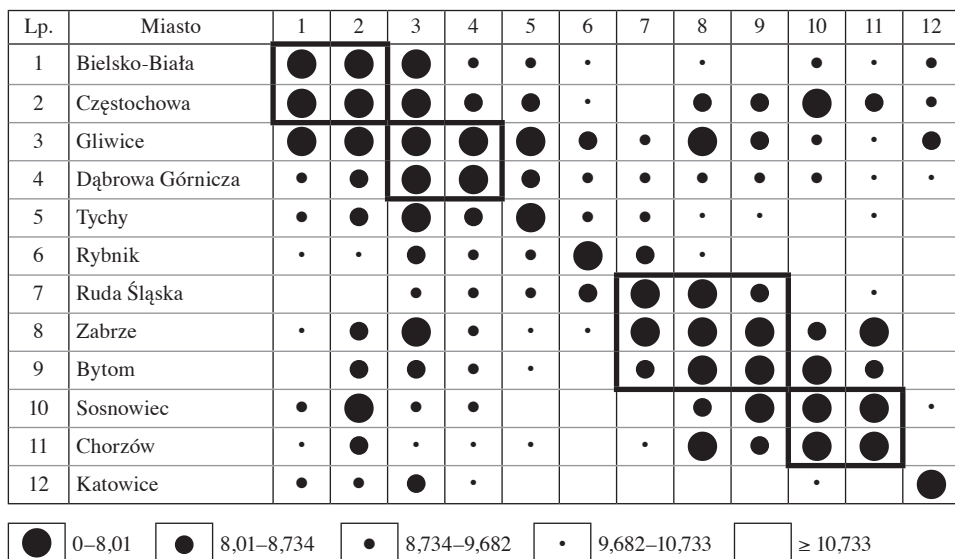
III – zmienne charakteryzujące poziom zamożności mieszkańców badanych miast: X_3 – X_8 , X_{12} – X_{14} , X_{29} – X_{30} ,

IV – zmienne charakteryzujące poziom kultury i edukacji: X_{21} – X_{28} .

3. Podobieństwo miast w 2002 roku

Rys. 1–4 przedstawiają uporządkowane diagramy Czekanowskiego dla określonych wyżej zbiorów zmiennych diagnostycznych dla 2002 r. Biorąc pod uwagę cały zbiór zmiennych diagnostycznych, wyróżniono siedem grup miast: 1) Bielsko-Biała i Częstochowa, 2) Dąbrowa Górnicza i Gliwice, 3) Ruda Śląska, Zabrze i Bytom, 4) Sosnowiec i Chorzów (rys. 1). Pozostałe trzy miasta, tj. Katowice, Rybnik i Tychy, nie wykazują wystarczająco bliskiego podobieństwa z innymi miastami i tworzą grupy jednoelementowe.

Na podstawie macierzy odległości euklidesowych można stwierdzić, że najmniejsze różnice pod względem wartości branych pod uwagę zmiennych opisujących sytuację demograficzno-społeczną miast w 2002 r. dzielą Bielsko-Białą i Częstochowę, a największe – Katowice i Rudę Śląską.

Rys. 1. Uporządkowany diagram Czekanowskiego dla zmiennych $X_{12}-X_{402}$, 2002 r.

Źródło: opracowanie własne.

Tabela 2. Porównanie średnich grupowych ze średnią ogólną, zmienne $X_{12}-X_{402}$, 2002 r.

Grupa/ zmienna	Iloraz średnich dla zmiennej													
	X_{12}	X_{22}	X_{32}	X_{42}	X_{52}	X_{62}	X_{72}	X_{82}	X_{92}	X_{102}	X_{112}	X_{122}	X_{132}	X_{142}
1	1,091	1,019	0,861	0,923	0,951	1,286	1,971	0,888	1,096	0,350	0,543	1,105	1,499	2,759
2	1,082	0,939	1,045	1,083	1,083	1,013	0,883	1,588	0,940	0,783	1,307	0,946	1,013	0,832
3	0,806	1,028	0,999	0,990	0,957	0,718	0,427	0,706	0,976	0,823	0,816	1,027	0,789	0,409
4	0,818	1,476	0,823	0,860	0,852	0,999	0,487	0,551	1,084	0,873	1,031	0,968	0,935	0,263
Grupa/ zmienna	Iloraz średnich dla zmiennej													
	X_{152}	X_{162}	X_{172}	X_{182}	X_{192}	X_{202}	X_{212}	X_{222}	X_{232}	X_{242}	X_{252}	X_{262}	X_{272}	X_{282}
1	0,770	0,660	0,897	0,770	0,961	0,961	0,880	1,034	1,092	0,938	1,200	1,044	1,532	1,319
2	1,167	0,658	1,338	0,893	0,652	0,765	0,854	1,036	1,025	1,331	0,943	0,783	0,971	0,943
3	1,189	1,008	1,073	1,241	1,005	1,229	1,272	0,617	0,689	1,064	1,387	1,031	0,584	0,951
4	0,977	1,512	0,914	0,799	1,321	1,243	1,038	0,943	1,013	1,299	0,588	1,313	0,985	0,650
Grupa/ zmienna	Iloraz średnich dla zmiennej													
	X_{292}	X_{302}	X_{312}	X_{322}	X_{332}	X_{342}	X_{352}	X_{362}	X_{372}	X_{382}	X_{392}	X_{402}		
1	1,059	0,917	0,784	0,988	1,035	1,088	1,077	1,150	1,014	0,890	0,934	0,425		
2	1,069	0,765	0,580	0,945	0,989	1,122	0,978	0,957	0,923	0,929	0,932	0,425		
3	0,882	0,886	1,249	1,063	0,980	0,744	0,938	0,777	1,026	0,987	1,087	1,640		
4	0,940	1,470	1,566	0,924	1,017	2,347	0,946	1,173	0,920	0,851	1,078	1,301		

Źródło: obliczenia własne.

Ze względu na wszystkie cechy najkorzystniejsza sytuacja w 2002 r. wystąpiła (tabela 2) w grupie pierwszej (Bielsko-Biała i Częstochowa). Znacznie wyższe niż średnia dla 12 miast były tu: liczba mieszkań oddanych do użytku na 1000 ludności (X_{142}), nakłady inwestycyjne na jednego mieszkańca w zł (X_{72}) oraz liczba uczniów liceów ogólnokształcących na 1000 ludności (X_{272}). O korzystnej sytuacji tej grupy miast świadczą także zmienne: emisja zanieczyszczeń gazowych (X_{102}) i emisja zanieczyszczeń pyłowych (X_{112}).

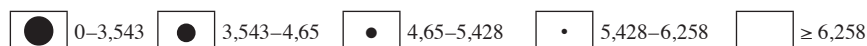
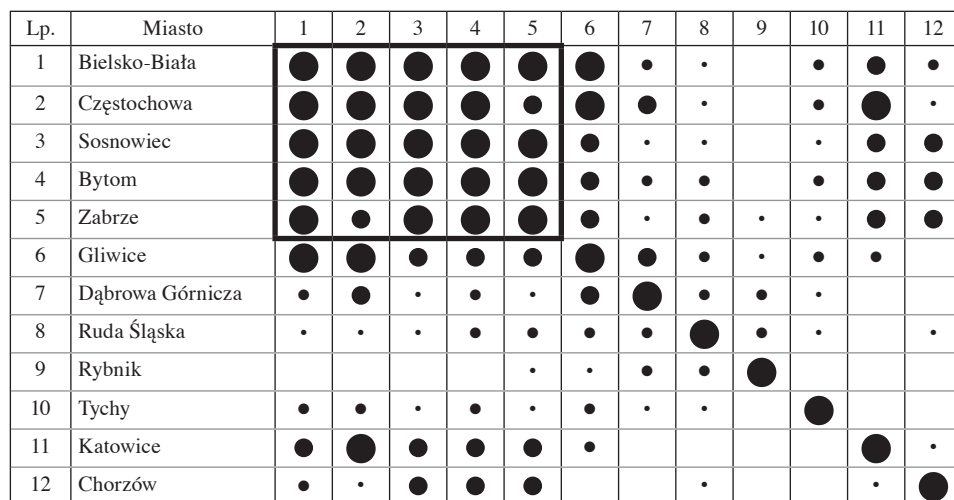
W drugiej grupie miast (Dąbrowa Górnicza i Gliwice) powyżej średniej dla 12 miast kształtowała się liczba ludności na jednego lekarza (X_{152}), liczba ludności na jedną pielęgniarkę (X_{172}) oraz liczba ludności na jedno miejsce w kinach stałych (X_{242}). Powyżej średniej dla 12 miast kształtowała się natomiast wartość brutto środków trwałych na jednego mieszkańca (X_{82}) i wartość współczynnika przyrostu naturalnego (X_{342}).

W najbardziej niekorzystnej sytuacji znajdowały się miasta grupy trzeciej (Ruda Śląska, Zabrze i Bytom), dla których większość badanych zmiennych przybierała zdecydowanie niekorzystne wartości, i grupy IV (Sosnowiec i Chorzów), w której wystąpiła znacznie wyższa od średniej dla 12 miast liczba bezrobotnych na 1000 pracujących (X_{22}), wyższa liczba ludności na jednego dentystę (X_{16}) i większa gęstość zaludnienia (X_{31}).

Ze względu na poziom zdrowotności i niektóre czynniki wpływające na jej poziom ($X_{152}-X_{202}$, $X_{92}-X_{112}$, X_{312} , $X_{382}-X_{392}$) wyróżniono tylko jedną grupę miast, do której należą (zob. rys. 2): Bielsko-Biała, Częstochowa, Sosnowiec, Zabrze i Bytom, przy czym najmniejsze, pod względem wartości branych pod uwagę współczynników, różnice dzieliły w 2002 r. Bielsko-Białą i Częstochowę, a największe – Katowice i Rybnik (tabela 3). Wśród miast niewykazujących podobieństwa do innych badanych miast województwa śląskiego (Gliwice, Dąbrowa Górnicza, Ruda Śląska, Rybnik, Tychy, Katowice i Chorzów), które tworzyły grupy jednoelementowe, najgorszą sytuację ze względu na zdrowie, a w szczególności jego uwarunkowania, zaobserwowano w Rudzie Śląskiej, a najkorzystniejszą w Katowicach.

Badając poziom zamożności (zmienne $X_{32}-X_{82}$, $X_{122}-X_{142}$, $X_{292}-X_{302}$) wyróżniono grupy: 1) Sosnowiec i Częstochowa, 2) Bytom i Zabrze, 3) Rybnik i Ruda Śląska, 4) Katowice i Gliwice oraz grupy jednoelementowe: Bielsko-Biała, Dąbrowa Górnicza, Tychy i Chorzów (rys. 3). Najmniejsze pod tym względem różnice dzieliły w 2002 r. Bytom i Zabrze, największe – Chorzów i Bielsko-Białą. Można dostrzec (tabela 4), że w tym przypadku najkorzystniejsza sytuacja wystąpiła w grupie czwartej, do której należą Katowice i Gliwice – wszystkie cechy diagnostyczne (z wyjątkiem liczby mieszkań oddanych do użytku na 1000 ludności) miały tu wartości wyższe niż średnia dla 12 miast. Poniżej średniej dla 12 miast kształtowała się większość badanych cech dla grupy drugiej (Bytom i Zabrze), wyjątek stanowiła zmienna X_{122} , i dla grupy trzeciej (Rybnik i Ruda

Śląska), wyjątkiem w tym przypadku były zmienne X_{32} i X_{52} . Ponadto znacznie poniżej średniej dla 12 miast dla grupy drugiej miast kształtowały się zmienne: nakłady inwestycyjne na jednego mieszkańca (X_{72}) i mieszkania oddane do użytku na 1000 ludności (X_{142}), co świadczy o niekorzystnej pod tym względem sytuacji. Najkorzystniejsza sytuacja ze względu na cechy charakteryzujące poziom zamożności spośród miast niewykazujących podobieństwa do innych badanych była w Tychach, a najmniej korzystna w Chorzowie.



Rys. 2. Uporządkowany diagram Czekanowskiego dla zmiennych charakteryzujących poziom zdrowotności $X_{152}-X_{202}$, $X_{92}-X_{112}$, X_{312} , $X_{382}-X_{392}$, 2002 r.

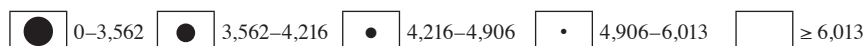
Źródło: opracowanie własne.

Tabela 3. Porównanie średnich grupowych ze średnią ogólną, zmienne opisujące poziom zdrowotności i niektóre jego uwarunkowania ($X_{92}-X_{112}$, $X_{152}-X_{202}$, X_{312} , $X_{382}-X_{392}$), 2002 r.

Grupa/ zmienna	Iloraz średnich dla zmiennej											
	X_{92}	X_{102}	X_{112}	X_{152}	X_{162}	X_{172}	X_{182}	X_{192}	X_{312}	X_{202}	X_{382}	X_{392}
1	1,055	0,406	0,498	0,939	0,778	0,893	0,914	1,101	1,129	1,141	0,835	0,979
Inne miasta												
Katowice	0,981	0,676	0,604	0,393	0,375	0,501	0,787	1,598	1,040	1,043	0,967	1,037
Ruda Śląska	0,913	1,246	1,114	1,361	1,485	1,382	1,480	0,613	1,006	0,974	1,151	1,241

Źródło: obliczenia własne.

Lp.	Miasto	1	2	3	4	5	6	7	8	9	10	11	12
1	Bielsko-Biała	●	●							•		•	
2	Częstochowa	●	●	●	●	●	•	•	•	•		•	•
3	Sosnowiec		●	●	●	●	•	•	•	•		•	●
4	Bytom		●	●	●	●	●	•	•	•		•	●
5	Zabrze		●	●	●	●	●	•	•	•	•	•	•
6	Ruda Śląska		•	•	●	●	●	●	•	•		●	•
7	Rybnik		•	•	•	•	●	●	●	•	•	•	•
8	Dąbrowa Górnicza		•	•	•	•	•	●	●	•	•	•	•
9	Gliwice	•	•	•	•	•	•	•	•	●	●	●	•
10	Katowice					•		•	•	●	●	•	
11	Tychy	•	•	•	•	•	●	•	•	●	•	●	•
12	Chorzów		•	●	●	•	•	•	•	•		•	●



Rys. 3. Uporządkowany diagram Czekanowskiego dla zmiennych charakteryzujących poziom zamożności $X_{32}-X_{82}$, $X_{122}-X_{142}$, $X_{292}-X_{302}$, 2002 r.

Źródło: opracowanie własne.

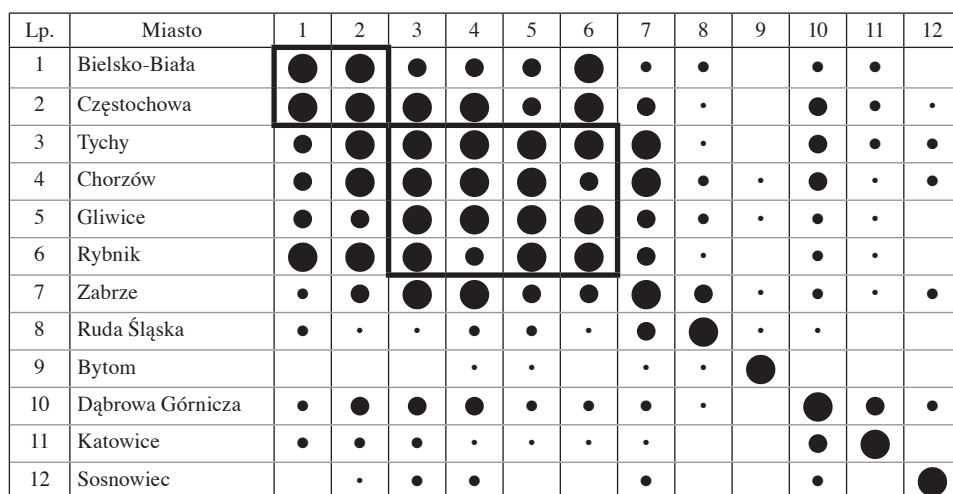
Tabela 4. Porównanie średnich grupowych ze średnią ogólną, zmienne charakteryzujące poziom zamożności $X_{32}-X_{82}$, $X_{122}-X_{142}$, $X_{292}-X_{302}$, 2002 r.

Grupa/zmienna	Iloraz średnich dla zmiennej										
	X_{32}	X_{42}	X_{52}	X_{62}	X_{72}	X_{82}	X_{122}	X_{132}	X_{142}	X_{292}	X_{302}
1	0,837	0,845	0,865	1,120	0,784	0,603	1,088	1,120	1,095	1,046	1,075
2	0,918	0,998	0,933	0,766	0,329	0,612	1,089	0,833	0,350	0,876	0,839
3	1,173	0,995	1,036	0,766	0,620	0,971	0,831	0,679	0,788	0,994	0,836
4	1,120	1,241	1,194	1,174	1,207	1,421	1,091	1,085	0,964	1,144	1,153
Inne miasta											
Chorzów	0,799	0,873	0,866	0,876	0,630	0,591	0,921	0,991	0,175	0,783	1,681
Tychy	1,219	0,983	1,031	1,026	2,087	1,268	0,959	1,105	1,226	0,970	1,023

Źródło: obliczenia własne.

Ze względu na poziom kultury i edukacji (zmienne $X_{232}-X_{282}$) można wśród badanych miast województwa śląskiego wyróżnić osiem grup miast: 1) Bielsko-Biała i Częstochowa, 2) Tychy, Gliwice, Chorzów i Rybnik oraz sześć grup

jednoelementowych: Zabrze, Ruda Śląska, Bytom, Dąbrowa Górnicza, Katowice i Sosnowiec, przy czym najmniejsze różnice dzieliły Chorzów i Tychy, a największe – Katowice i Bytom (rys. 4). Lepszą pod tym względem sytuację odnotowano w Bielsku-Białej i Częstochowie (grupa pierwsza), dla których prawie wszystkie zmienne miały wartości wyższe niż średnia dla 12 miast (tabela 5), wyjątek stanowiła zmienna X_{242} – liczba ludności na jedno miejsce w kinach stałych. Natomiast w grupie drugiej powyżej średniej dla 12 miast kształtowały się cechy świadczące o poziomie edukacji: liczba uczniów liceów ogólnokształcących i uczniów ponadgimnazjalnych zasadniczych szkół zawodowych na 1000 ludności.



Rys. 4. Uporządkowany diagram Czekanowskiego dla zmiennych charakteryzujących poziom kultury i edukacji X_{232} – X_{282} , 2002 r.

Źródło: opracowanie własne.

Tabela 5. Porównanie średnich grupowych ze średnią ogólną, zmienne charakteryzujące poziom kultury i edukacji X_{232} – X_{282} , 2002 r.

Grupa/zmienna	Iloraz średnich dla zmiennej					
	X_{232}	X_{242}	X_{252}	X_{262}	X_{272}	X_{282}
1	1,092	0,938	1,200	1,044	1,532	1,319
2	0,992	0,779	0,595	0,832	1,097	1,185

Źródło: obliczenia własne.

4. Podobieństwo miast w 2009 roku

Ze względu na wszystkie zmienne diagnostyczne (zmienne X_{19} – X_{409}) dla 2009 r. wyróżniono grupy miast: 1) Bielsko-Biała i Częstochowa, 2) Sosnowiec, Zabrze i Gliwice oraz grupy jednoelementowe: Bytom, Ruda Śląska, Chorzów, Rybnik, Tychy, Dąbrowa Górnicza i Katowice (rys. 5).

Najmniejsze różnice dzieliły w 2009 r. Bielsko-Białą i Częstochowę, a największe – Katowice i Rybnik. Bardziej niekorzystna sytuacja ze względu na cechy X_{19} – X_{409} (tabela 6) wystąpiła w grupie drugiej (Sosnowiec, Zabrze i Gliwice). Znacznie wyższe niż średnia dla 12 miast okazały się w tym przypadku: współczynnik zgonów niemowląt (X_{209}), liczba bezrobotnych na 1000 pracujących (X_{29}), gęstość zaludnienia (X_{319}), liczba ludności na jednego lekarza i na jedną pielęgniarkę (X_{159} i X_{179}), na korzyść tej grupy miast świadczy natomiast wartość współczynnika przyrostu naturalnego (X_{349}).

W grupie pierwszej (Bielsko-Biała i Częstochowa) powyżej średniej dla 12 miast kształtowały się: liczba mieszkań oddanych do użytkowania na 1000 ludności (X_{149}), zużycie gazu w gospodarstwach domowych na jednego mieszkańca (X_{139}), liczba podmiotów zarejestrowanych w rejestrze Regon na 10 tys. ludności (X_{69}), liczba uczniów liceów ogólnokształcących i ponadgimnazjalnych zasadniczych szkół zawodowych (X_{279} i X_{289}), liczba pracujących ogółem na 1000 ludności (X_{19}) – zmienne świadczące o poziomie zamożności i poziomie edukacji.

Lp.	Miasto	1	2	3	4	5	6	7	8	9	10	11	12
1	Bielsko-Biała	●	●	•	●	•			•	•	●	•	●
2	Częstochowa	●	●	●	●	•	•	•	•		●	•	•
3	Sosnowiec	•	●	●	●	●	●	•	•			•	•
4	Gliwice	●	●	●	●	●	•	•	•	•	●	●	•
5	Zabrze	•	•	●	●	●	•	•	•	•	•	•	
6	Bytom		•	●	•	•	●	•	•				
7	Ruda Śląska		•	•	•	•	•	●	•	•	•	•	
8	Chorzów	•	•	•	•	•	•	•	●	•	•	•	
9	Rybnik	•			•	•		•	•	●	•	•	
10	Tychy	●	●		●	•		•	•	•	●	•	•
11	Dąbrowa Górnicza	•	•	•	●	•		•	•	•	•	●	
12	Katowice	●	•	•	•						•		●

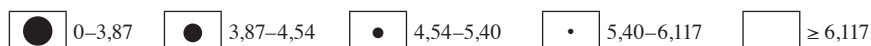
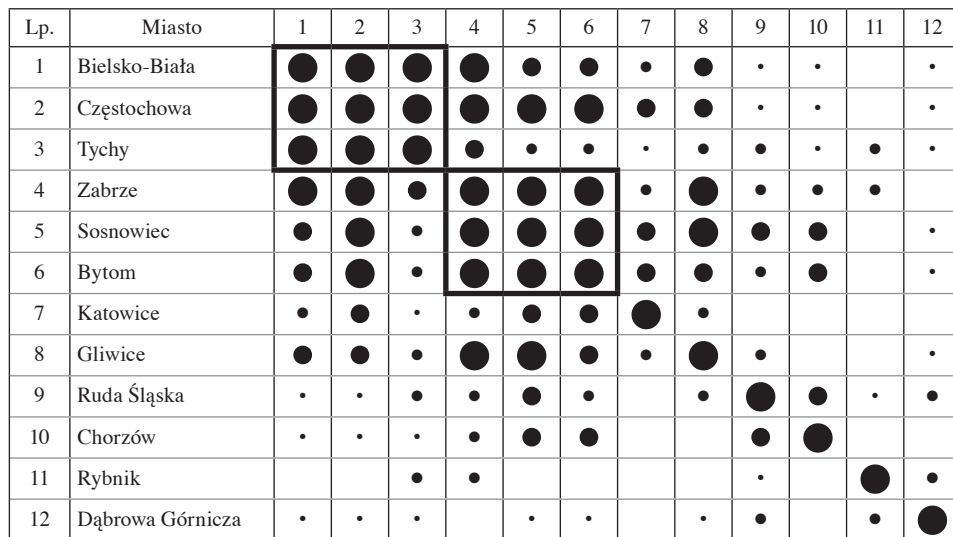


Rys. 5. Uporządkowany diagram Czekanowskiego dla wszystkich zmiennych X_{19} – X_{409} , 2009 r.
Źródło: opracowanie własne.

Tabela 6. Porównanie średnich grupowych ze średnią ogólną, wszystkie zmienne diagnostyczne X_{19} – X_{409} , 2009 r.

Grupa/ zmienna	Iloraz średnich dla zmiennej													
	X_{19}	X_{29}	X_{39}	X_{49}	X_{59}	X_{69}	X_{79}	X_{89}	X_{99}	X_{109}	X_{119}	X_{129}	X_{139}	X_{149}
1	1,157	0,854	0,945	1,002	0,967	1,201	1,126	0,874	1,076	0,312	0,700	1,133	1,568	1,623
2	0,884	1,160	1,009	1,014	1,015	0,999	0,892	0,872	0,915	0,259	0,569	1,003	0,948	0,918
Grupa/ zmienna	Iloraz średnich dla zmiennej													
	X_{159}	X_{169}	X_{179}	X_{189}	X_{199}	X_{209}	X_{219}	X_{229}	X_{239}	X_{249}	X_{259}	X_{269}	X_{279}	X_{289}
1	0,750	0,742	0,873	0,708	1,060	0,696	0,859	0,950	1,221	0,807	1,036	0,845	1,617	1,340
2	1,017	0,994	1,043	0,949	0,941	1,396	1,028	0,836	0,967	0,698	0,831	1,146	0,819	0,884
Grupa/ zmienna	Iloraz średnich dla zmiennej													
	X_{299}	X_{309}	X_{319}	X_{329}	X_{339}	X_{349}	X_{359}	X_{369}	X_{379}	X_{389}	X_{399}	X_{409}		
1	1,051	0,906	0,790	0,999	1,035	0,362	0,997	1,147	1,024	0,940	0,922	0,961		
2	1,009	1,017	1,124	0,952	1,000	2,303	0,951	1,002	0,927	0,951	0,967	1,219		

Źródło: obliczenia własne.

Rys. 6. Uporządkowany diagram Czekanowskiego dla zmiennych charakteryzujących poziom zdrowotności X_{99} – X_{119} , X_{159} – X_{209} , X_{319} , X_{389} – X_{399} , 2009 r.

Źródło: opracowanie własne.

Ze względu na zmienne charakteryzujące poziom zdrowotności i jego niektóre uwarunkowania w 2009 r. wyróżniono grupy miast: 1) Bielsko-Biała, Częstochowa i Tychy, 2) Sosnowiec, Zabrze i Bytom oraz grupy jednoelementowe: Katowice, Gliwice, Ruda Śląska, Chorzów, Rybnik i Dąbrowa Górnicza (rys. 6). Najmniejsze różnice dzieliły Bielsko-Białą i Częstochowę, a największe – Katowice i Rybnik. Korzystna sytuacja (zob. tabela 8) wystąpiła w grupie pierwszej (Bielsko-Biała, Częstochowa i Tychy). Wyższy od średniej dla 12 miast był współczynnik ścieków oczyszczonych (X_{92}) oraz liczba łóżek w szpitalach na 10 tys. ludności (X_{192}), pozostałe zmienne przybrały wartości poniżej średniej dla 12 miast. Miasta grupy drugiej (Sosnowiec, Zabrze i Bytom) zdecydowanie niekorzystnie wypadły natomiast pod względem wartości współczynnika zgonów niemowląt, gęstości zaludnienia i standaryzowanego współczynnika zgonów kobiet.

Tabela 7. Porównanie średnich grupowych ze średnią ogólną, zmienne charakteryzujące poziom zdrowotności X_{99} – X_{119} , X_{159} – X_{209} , X_{319} , X_{389} – X_{399} , 2009 r.

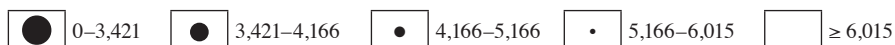
Grupa/zmienna	Iloraz średnich dla zmiennej											
	X_{99}	X_{109}	X_{119}	X_{159}	X_{169}	X_{179}	X_{189}	X_{199}	X_{319}	X_{209}	X_{389}	X_{399}
1	1,078	0,340	0,581	0,815	0,878	0,917	0,851	1,043	0,814	0,565	0,925	0,897
2	0,989	0,320	0,869	0,978	0,935	0,928	0,967	1,140	1,336	1,372	0,994	1,025
Inne miasta												
Katowice	0,800	0,514	0,771	0,451	0,685	0,471	0,758	1,661	1,018	1,162	1,007	1,023
Ruda Śląska	0,974	0,498	0,678	1,094	1,302	1,491	1,258	0,530	1,003	0,798	1,208	1,138
Rybnik	1,074	3,843	1,539	1,605	1,589	0,809	1,579	0,784	0,518	1,022	0,908	0,917
Dąbrowa Górnicza	1,077	2,368	3,365	1,394	0,855	1,331	0,896	0,585	0,368	0,812	1,046	1,087

Źródło: obliczenia własne.

Ze względu na poziom zamożności (zmienne X_{39} – X_{89} , X_{129} – X_{149} , X_{299} – X_{309}) można wśród badanych miast województwa śląskiego wyróżnić grupy: 1) Dąbrowa Górnicza i Gliwice, 2) Rybnik, Ruda Śląska, Bytom i Chorzów, 3) Sosnowiec i Częstochowa oraz grupy jednoelementowe: Bielsko-Biała, Tychy, Katowice i Zabrze, przy czym najmniejsze różnice dzieliły Bytom i Rudę Śląską, a największe – Katowice i Bytom (rys. 7).

Najkorzystniejsza sytuacja pod względem zamożności wystąpiła w pierwszej grupie miast (Gliwice i Dąbrowa Górnicza); prawie wszystkie cechy miały wartości wyższe niż średnia dla 12 miast, wyjątek stanowią cechy: X_{59} – wydatki budżetu miasta na jednego mieszkańca, X_{129} – zużycie wody w gospodarstwie domowym na jednego mieszkańca, oraz X_{309} – drogi publiczne powiatowe o twardej nawierzchni (tabela 8).

Lp.	Miasto	1	2	3	4	5	6	7	8	9	10	11	12
1	Bielsko-Biała	●	●	•	•				•	●	•	●	
2	Gliwice	●	●	●	•	•	•	•	•	●	●	●	•
3	Dąbrowa Górnicza	•	●	●	●	•	•	•	•	•	•	•	•
4	Rybnik	•	•	●	●	●	●	•	•	●	•		●
5	Ruda Śląska		•	•	●	●	●	●	•	•			●
6	Bytom		•	•	●	●	●	●	●	•			●
7	Chorzów		•	•	•	●	●	●	●	•			•
8	Sosnowiec	•	•	•	•	•	●	●	●	●			•
9	Częstochowa	●	●	•	•	•	•	•	●	●	•	•	•
10	Tychy	•	●	•	•					•	●	•	•
11	Katowice	●	●	•						•	•	●	
12	Zabrze		•	•	•	•	•	•	•	•	•		●



Rys. 7. Uporządkowany diagram Czekanowskiego dla zmiennych charakteryzujących poziom zamożności X_{39} – X_{89} , X_{129} – X_{149} , X_{299} – X_{309} , 2009 r.

Źródło: opracowanie własne.

Tabela 8. Porównanie średnich grupowych ze średnią ogólną, zmienne charakteryzujące poziom zamożności X_{39} – X_{89} , X_{129} – X_{149} , X_{299} – X_{309} , 2009 r.

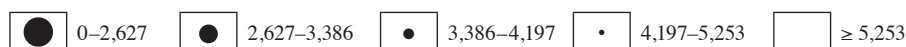
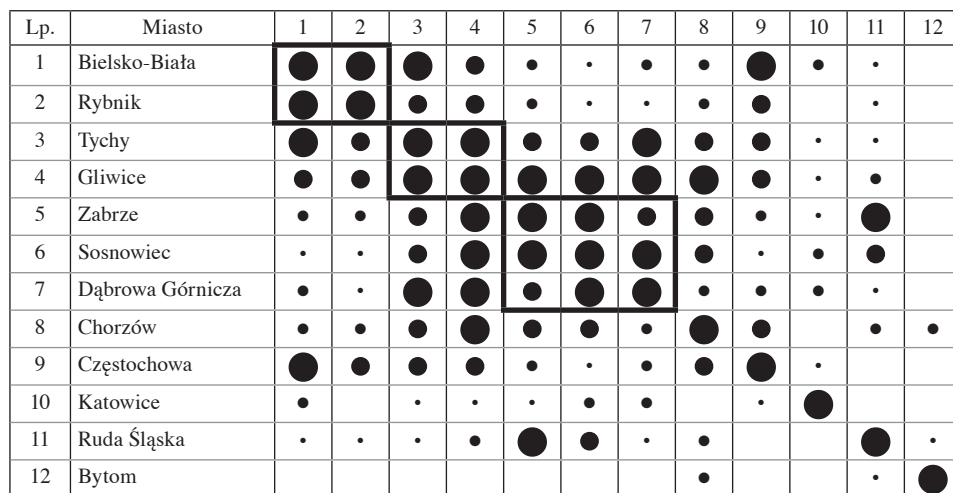
Grupa/zmienna	Iloraz średnich dla zmiennej										
	X_{39}	X_{49}	X_{59}	X_{69}	X_{79}	X_{89}	X_{129}	X_{139}	X_{149}	X_{299}	X_{309}
1	1,120	1,028	0,944	1,028	1,307	1,627	0,947	1,000	1,254	1,094	0,663
2	0,906	0,892	0,901	0,843	0,610	0,628	0,895	0,778	0,504	0,898	1,058
3	0,903	0,822	0,815	1,056	0,573	0,584	1,148	0,994	1,131	1,018	1,206
Inne miasta											
Bielsko-Biała	1,001	1,112	1,220	1,343	1,487	1,134	1,087	1,931	1,770	1,092	0,953
Zabrze	0,991	1,238	1,438	0,828	0,692	0,659	0,889	0,945	0,197	0,869	0,691
Katowice	1,378	1,180	1,278	1,303	1,573	1,527	1,144	1,133	1,525	1,205	1,335

Źródło: obliczenia własne.

Poniżej średniej dla 12 miast (sytuacja niekorzystna) kształtowała się większość badanych cech dla grupy drugiej (Rybnik, Ruda Śląska, Bytom, Chorzów i Sosnowiec), wyjątek stanowiła tu zmienna X_{309} – drogi publiczne powiatowe o twardej nawierzchni. Spośród miast, które nie zostały przydzielone do żadnej grupy, najkorzystniejszą sytuację pod względem poziomu zamożności (i bardziej

korzystną niż w grupie pierwszej) zaobserwowano w Katowicach i Bielsku-Białej, a najgorszą w Zabrze.

Rozpatrując poziom kultury i edukacji ($X_{239}-X_{289}$) można wyróżnić grupy miast: 1) Bielsko-Biała i Rybnik, 2) Tychy i Gliwice, 3) Zabrze, Sosnowiec i Dąbrowa Górnicza oraz grupy jednoelementowe: Chorzów, Częstochowa, Katowice, Ruda Śląska i Bytom, przy czym najmniejsze różnice dzieliły Gliwice i Zabrze, a największe – Katowice i Bytom (rys. 8).



Rys. 8. Uporządkowany diagram Czekanowskiego dla zmiennych charakteryzujących poziom kultury i edukacji $X_{239}-X_{289}$, 2009 r.

Źródło: opracowanie własne.

Tabela 9. Porównanie średnich grupowych ze średnią ogólną, zmienne charakteryzujące poziom kultury i edukacji $X_{239}-X_{289}$, 2009 r.

Grupa/zmienna	Iloraz średnich dla zmiennej					
	X_{239}	X_{249}	X_{259}	X_{269}	X_{279}	X_{289}
1	1,215	0,661	0,984	0,797	1,194	1,688
2	1,006	0,798	0,812	0,884	0,954	1,047
3	1,003	0,718	0,862	1,124	0,808	0,610
Inne miasta						
Bytom	0,346	3,288	1,497	1,030	0,654	0,847
Częstochowa	1,048	0,665	0,953	0,893	1,884	1,162

Źródło: obliczenia własne.

W pierwszej grupie (Bielsko-Biała i Rybnik) poziom trzech zmiennych (tj. przypadającej na 1000 ludności liczby czytelników – X_{239} , liczby uczniów liceów ogólnokształcących – X_{279} , i uczniów ponadgimnazjalnych zasadniczych szkół zawodowych – X_{289}) był zdecydowanie wyższy niż średnia dla 12 miast (tabela 9). W grupie trzeciej (Zabrze, Sosnowiec i Dąbrowa Górnicza) przypada zdecydowanie więcej uczniów gimnazjum na jeden komputer z dostępem do internetu niż średnio dla 12 miast. Wśród miast nienależących do żadnej z grup najlepsza sytuacja pod względem poziomu kultury i edukacji była w Częstochowie, a najgorsza w Bytomiu.

5. Podsumowanie

Na podstawie przeprowadzonej analizy można uznać, że najbardziej podobnymi do siebie, a także najlepszymi pod względem sytuacji społeczno-gospodarczej i demograficznej były zarówno w 2002 r., jak i w 2009 r., mimo dzielącej je sporej odległości geograficznej, Bielsko-Biała i Częstochowa – miasta leżące poza obszarem byłego Górnośląskiego Okręgu Przemysłowego, na terenach zielonych województwa śląskiego. Cechowały się one najkorzystniejszymi na ogół wartościami całego zbioru zmiennych diagnostycznych, jak również zbiorów obejmujących zmienne charakteryzujące sytuację zdrowotną miast i jej uwarunkowania oraz poziomu kultury i edukacji. Miasto Bielsko-Biała cechowało się ponadto wysokim poziomem zamożności, nie wykazując jednak w tym względzie podobieństwa do innych miast.

Najbardziej podobnymi do siebie, a zarazem będącymi w niekorzystnej sytuacji miastami były: Sosnowiec, Zabrze i Gliwice – miasta leżące na terenie byłego Górnośląskiego Okręgu Przemysłowego.

Stosując metodę Czekanowskiego do badania zagadnień społeczno-demograficznych, warto zwrócić także uwagę na obiekty niepodobne do innych, co może okazać się ważne dla wyjaśniania przyczyn podobieństwa i odmienności badanych miast.

Literatura

- Czekanowski J. [1909], *Zur Differentialdiagnose der Neandertalgruppe*, Korespondenz-Blatt der deutschen Gesellschaft für Anthropologie, Ethnologie und Urgeschichte, Braunschweig, vol. 40.
- Czekanowski J. [1913], *Zarys metod statystycznych w zastosowaniu do antropologii*, Prace Naukowego Towarzystwa Warszawskiego, nr 5.
- Heffner K., Gibas P. [2007], *Analiza ekonomiczno-przestrzenna*, Wydawnictwo Akademii Ekonomicznej w Katowicach, Katowice.

- Lipieta A. i in. [2000], *Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.
- Mielecka-Kubień Z., Warzecha K. [2010], *Poziom rozwoju społeczno-gospodarczego dużych miast woj. śląskiego i jego uwarunkowania*, Sprawozdanie z Badań Statutowych Katedry Ekonometrii Uniwersytetu Ekonomicznego w Katowicach (maszynopis).
- Pociecha J. i in. [1988], *Metody taksonomiczne w badaniach społeczno-ekonomicznych*, PWN, Warszawa.
- Pluta W. [1986], *Wielowymiarowa analiza porównawcza w modelowaniu ekonometrycznym*, PWN, Warszawa.
- Wolny R. [2009], *Metody ilościowe w badaniach rynku*, Wydawnictwo Akademii Ekonomicznej w Katowicach, Katowice.

The Socio-Demographic Situation of Chosen Cities in Śląskie Province in the Years 2002 and 2009

The purpose of the research was to determine the similarities that exist among cities in the province of Śląskie with more than 100,000 inhabitants in the interval of a few years. The cities considered were Bielsko-Biała, Bytom, Chorzów, Częstochowa, Dąbrowa Górnicza, Gliwice, Katowice, Ruda Śląska, Rybnik, Sosnowiec, Tychy and Zabrze, all of which are located within close geographical proximity to one another.

The similarity of the cities was analysed according to chosen socio-economic indicators and demographic coefficients in the years 2002 and 2009. J. Czekanowski method was applied; to determine which groups of cities were better or worse off with regard to the given profile of the investigation, the average values of the indicators for the entire set of the cities, and for its established groups were compared.

The research was conducted in the following profiles: I – socio-demographic situation, II – population health and its determinants, III – affluence, IV – educational and cultural level. The cities that were the most similar to each other, and in the best socio-economic and demographic situation were, for both years, and in spite of the relatively big geographic distance, Bielsko-Biała and Częstochowa. Bielsko-Biała stood alone among the cities for its affluence. Sosnowiec, Zabrze and Gliwice – all cities in the former Upper-Silesian Industrial District – were similar to each other insofar as they were in the worst situation.

Keywords: J. Czekanowski's method, the socio-demographic situation, cities in Śląskie Province, taxonomic methods.

Marcin Salamaga

Katedra Statystyki

Uniwersytet Ekonomiczny w Krakowie

Propozycja modyfikacji wskaźnika ujawnionej przewagi komparatywnej z wykorzystaniem funkcji wymiernych

Streszczenie

Wskaźnik względnej przewagi komparatywnej (*RCA*) jest wykorzystywany powszechnie do oceny konkurencyjności eksportu towarów i usług oraz przewagi komparatywnej kraju według sektorów produkcji lub grup towarowych. Jego wartości nie są jednak unormowane, co skutkuje określonymi konsekwencjami. Rozkłady wskaźnika *RCA* są zwykle asymetryczne, niestabilne w czasie, wrażliwe na liczbę sektorów (grup towarowych) i na liczbę krajów w grupie referencyjnej. Utrudnia to porównywalność wartości wskaźnika w ujęciu czasowym i przestrzennym oraz komplikuje wyznaczenie odpowiedniego rozkładu teoretycznego. W artykule zaproponowano pewną metodę transformacji wskaźnika *RCA*, która umożliwi poprawę niektórych jego własności. Proponowana metoda wykorzystuje transformację przedmiotowego wskaźnika za pomocą rodziny funkcji wymiernych. W pracy analizowano empiryczne rozkłady unormowanego wskaźnika *RCA* dla towarów grupowanych zgodnie ze standardową międzynarodową klasyfikacją handlu (SITC).

Słowa kluczowe: wskaźnik względnej przewagi komparatywnej, eksport, rozkład stabilny, test chi-kwadrat.

1. Wprowadzenie

W badaniach nad konkurencyjnością eksportu towarów i usług oraz przewag komparatywnych kraju według sektorów produkcji bądź grup towarowych bardzo często wykorzystuje się wskaźnik względnej przewagi komparatywnej (*RCA*) zaproponowany przez B. Balassę [1965]. Wskaźnik ten jest zdefiniowany jako stosunek wartości eksportu badanej grupy towarowej analizowanego kraju w całkowitej wartości eksportu kraju do udziału wartości światowego eksportu wyróżnionej grupy towarów w całkowitej wartości światowego eksportu. Miernik Balassy powstał w ramach nurtu badań nad międzynarodowym konkurowaniem gospodarek, który został zainspirowany zasadą kosztów względnych D. Ricardo [1957]. Obecnie w licznych analizach stosuje się również reinterpretacje oryginalnej formuły Balassy, np. zastępując całkowitą wartość eksportu przez wartość importu odpowiednich towarów [Misala 2011].

Wskaźnik względnej przewagi komparatywnej cechuje się niewielką uciążliwością obliczeniową i łatwą interpretacją wyników. Ma on jednak kilka wad wynikających m.in. z braku unormowania jego wartości.

Empiryczne rozkłady sektorowego wskaźnika *RCA* cechują się na ogół silną asymetrią i brakiem stabilności w czasie, co w istotny sposób ogranicza porównywalność jego wartości w ujęciu czasowym i przestrzennym. Za kolejną wadę należy uznać to, że wartości wskaźnika są wrażliwe zarówno na liczbę sektorów (grup towarowych), jak i na liczbę krajów w grupie referencyjnej. Skutkuje to brakiem stabilności ważnych charakterystyk liczbowych miernika i stanowi utrudnienie przy wyznaczeniu teoretycznego rozkładu jego wartości. W literaturze przedmiotu podejmowane są próby konstrukcji unormowanych wskaźników względnej przewagi komparatywnej, np. addytywny wskaźnik *RCA* [Hoen i Oosterhaven 2006], znormalizowany indeks względnej przewagi komparatywnej [Yu, Cai i Leung 2000], wskaźnik względnej symetrycznej przewagi komparatywnej [Dalum, Laursen i Villumsen 1998] oraz [Iapadre 2001]. Nie zyskały one jednak do tej pory trwałej popularności w empirycznych zastosowaniach wśród badaczy międzynarodowej wymiany handlowej.

W opracowaniu zaproponowano metodę transformacji wskaźnika *RCA* o wartościach z przedziału $[0; \infty)$ do przedziału liczbowego od -1 do 1 . Zastosowano do tego celu pewną rodzinę funkcji wymiernych. Proponowane przekształcenie powinno poprawić niektóre własności miernika względnej przewagi komparatywnej. Empiryczne rozkłady zmodyfikowanego wskaźnika *RCA* badano na podstawie danych o międzynarodowej wymianie handlowej Polski. W obliczeniach wykorzystano dane Eurostatu z lat 2005–2009.

2. Konstrukcja skorygowanego wskaźnika ujawnionej przewagi komparatywnej

Podstawą konstrukcji skorygowanego miernika względnej przewagi komparatywnej jest wskaźnik Balassy [1965], który można wyrazić następującą formułą:

$$RCA_i = \frac{X_{ij}}{X_j} \cdot \frac{X_i}{X}, \quad (1)$$

gdzie:

X_{ij} – wartość eksportu i -tej grupy towarowej w j -tym kraju,

X_j – całkowita wartość eksportu j -tego kraju,

X_i – wartość światowego eksportu i -tej grupy towarowej,

X – całkowita wartość eksportu światowego.

Wartości wskaźnika (1) większe od jednościi wskazują na występowanie przewagi komparatywnej, a także określają jej intensywność. Wartości mniejsze od 1 wskazują na brak tego typu przewagi. Miernik (1) jest również stosowany w sytuacji oceny konkurencyjności eksportu kraju w porównaniu z krajami referencyjnymi (np. krajami regionu geograficznego, ugrupowaniem gospodarczym czy krajem partnerskim). Odpowiednia formuła ma wówczas postać:

$$RCA_i = \frac{X_{ij}}{X_j} \cdot \frac{X_i^R}{X^R}, \quad (2)$$

gdzie:

X_{ij} – wartość eksportu i -tej grupy towarowej w j -tym kraju,

X_j – całkowita wartość eksportu j -tego kraju,

X_i^R – wartość światowego eksportu i -tej grupy towarowej w krajach referencyjnych,

X^R – całkowita wartość eksportu światowego w krajach referencyjnych.

Fakt, że wskaźnik względnej przewagi komparatywnej nie jest unormowany, powoduje, iż jego rozkłady empiryczne zazwyczaj cechują się asymetrią i brakiem stabilności w ujęciu czasowym i przestrzennym. Może to sprawiać pewne trudności w porównywaniu rozkładów wskaźnika RCA obliczanego np. dla różnych grup towarowych.

W artykule proponuje się przekształcenie wskaźnika RCA , które poprawia niektóre z jego własności. Przekształcenie polega na wykorzystaniu klasy funkcji wymiernych o następującej postaci:

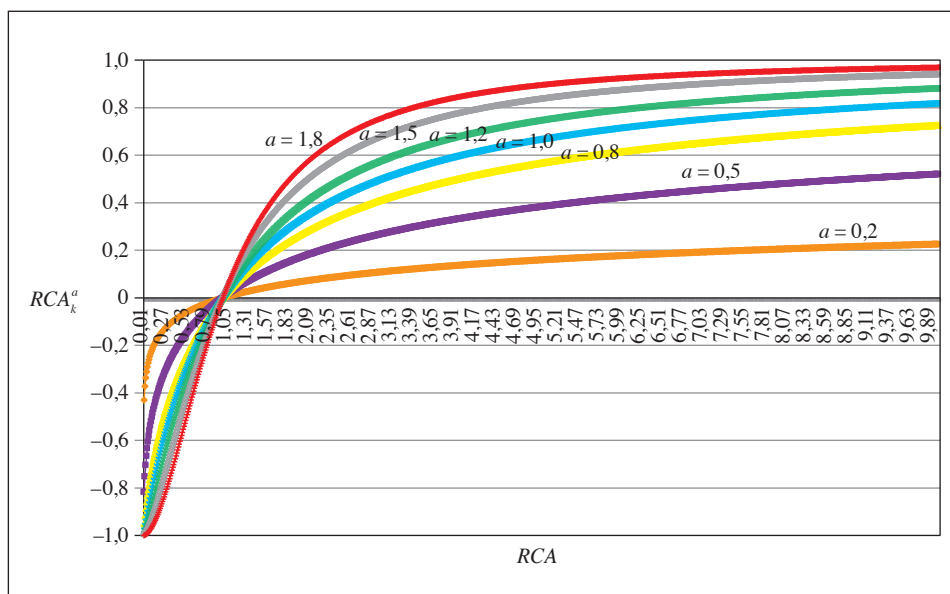
$$f(x) = \frac{x^a - 1}{x^a + 1}, \quad (3)$$

gdzie a to dowolna liczba dodatnia.

Proponowany skorygowany wskaźnik względnej przewagi komparatywnej ma więc postać:

$$RCA_k^a = \frac{RCA^a - 1}{RCA^a + 1}. \quad (4)$$

Konstrukcja wskaźnika (4) jest inspirowana miarą zaproponowaną przez B. Daluma, K. Laursena i G. Villumsena [1998] i stanowi jej uogólnienie. Przekształcenie (3) pozwala uzyskać unormowaną wartość wskaźnika RCA_k^a w przedziale od -1 do 1 . Dodatnie wartości wskaźnika (4) wskazują na występowanie ujawnionej przewagi komparatywnej w eksporcie, zaś ujemne wartości tego miernika wskazują na brak takiej przewagi. Od przyjętej wartości parametru a zależy tempo zbieżności wskaźnika RCA_k^a do wartości granicznych przedziału $[-1; 1)$. Na rys. 1 przedstawiono przebieg wartości miernika RCA_k^a w zależności od wartości oryginalnego wskaźnika RCA dla wybranych poziomów parametru a .



Rys. 1. Przebieg wartości wskaźnika RCA_k^a w zależności od wskaźnika RCA dla wybranych wartości parametru a

Źródło: opracowanie własne.

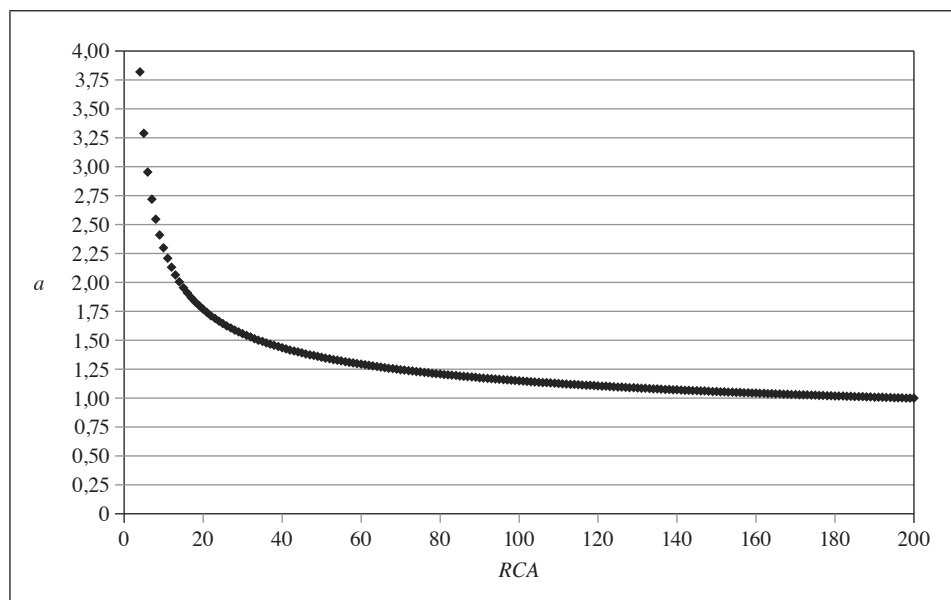
Im wyższy jest poziom parametru a , tym intensywniej wartości miernika RCA_k^a zbiegają do wartości granicznych przedziału zmodyfikowanego wskaźnika. Natomiast im poziom parametru a jest bliższy 0, tym wolniej wartości miernika RCA_k^a zbiegają do dolnej i górnej granicy przedziału $[-1; 1)$. W praktyce

wybór wartości parametru a decyduje, jakie będą rzeczywiste wartości ekstremalne przekształconego wskaźnika RCA oraz jakie jest tempo jego „zbliżania się” do wartości granicznych równych odpowiednio -1 i 1 .

Przyjmijmy przykładowo, że liczba $\pm 0,99$ jest granicą „dostatecznej bliskości” dla 1 i -1 . Wówczas wymaganą wartość parametru a gwarantującą spełnienie warunku $|RCA_k^a| > 0,99$ w zależności od maksymalnej wartości wskaźnika RCA można opisać zależnością:

$$a(RCA) = \begin{cases} \frac{\ln(199)}{\ln(RCA)} & \text{dla } RCA > 1 \\ -\frac{\ln(199)}{\ln(RCA)} & \text{dla } RCA \in (0;1) \end{cases} \quad (5)$$

W miarę wzrostu wartości wskaźnika (1) poziom parametru a zapewniający zbieżność skorygowanego wskaźnika RCA do górnego kresu zbioru $[-1, 1]$ z 1-procentową dokładnością maleje asymptotycznie do 0. Jeśli np. maksymalna wartość wskaźnika względnej przewagi komparatywnej wynosi 200, to konieczne jest przyjęcie wykładnika a równego ok. 1, aby uzyskać zbieżność do górnej granicy przedziału $[-1, 1]$ z wymaganą dokładnością (por. rys. 2).



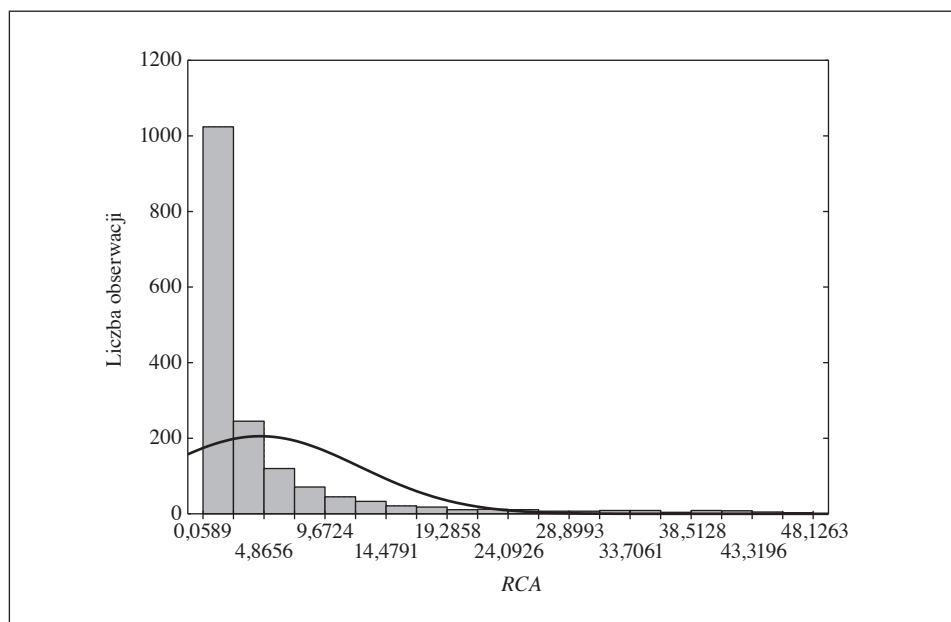
Rys. 2. Wykres funkcji $a(RCA)$ dla $RCA > 1$

Źródło: opracowanie własne.

Należy zauważyć, że przyjęcie zbyt niskiej wartości parametru a (przy ustalonym poziomie wskaźnika RCA) spowoduje oddolne i odgórne zawężenie obszaru zmienności wskaźnika RCA_k^a .

3. Ocena stabilności empirycznego rozkładu wartości wskaźnika przewagi komparatywnej w czasie

Empiryczne rozkłady wskaźnika ujawnionej przewagi komparatywnej analizowano na przykładzie Polski względem Słowacji jako kraju referencyjnego. Do zbadania stabilności empirycznych rozkładów wskaźników RCA i $RCA_k^{(a)}$ w czasie zastosowano test zgodności chi-kwadrat dla liczebności obserwowanych względem oczekiwanych (jako oczekiwane liczebności przyjęto rozkłady częstości wskaźnika z poprzednich lat w stosunku do wybranego roku badawczego).

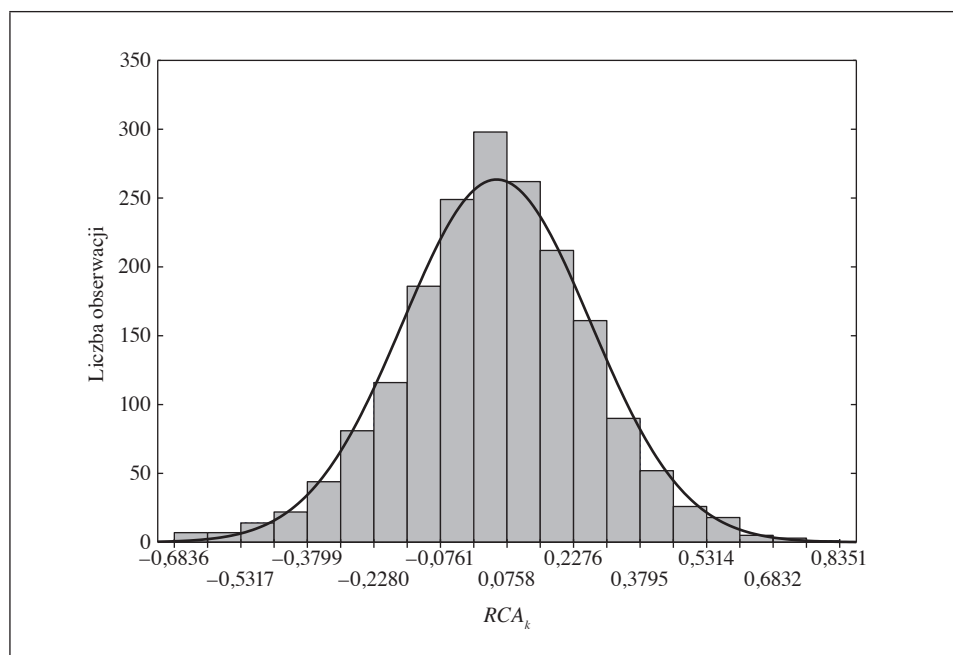


Rys. 3. Rozkład wartości wskaźnika RCA dla Polski względem Słowacji jako kraju referencyjnego w 2005 r.

Źródło: opracowanie własne na podstawie danych Eurostatu.

W tym celu wartości wskaźników RCA i $RCA_k^{(a)}$ w każdym z lat 2005–2006 zostały pogrupowane w szeregi rozdzielcze przedziałowe. W analizie posłużono się szeregami składającymi się z 20 klas, zachowując postulat podziału rozłącz-

nego i wyczerpującego [Sobczyk 2000]. Z większości przypadków tworzono szeregi o jednakowych interwałach klasowych z wyjątkiem sytuacji występowania obserwacji ekstremalnych – wówczas wprowadzano otwarte skrajne klasy w szeregu statystycznym. Na rys. 3 i 4 przedstawiono przykładowe rozkłady wartości wskaźników RCA i $RCA_k^{(a)}$ dla $a = 0,2$ obliczone dla Polski względem Słowacji jako kraju referencyjnego w 2005 r. Ze względu na znaczny zakres zmienności wskaźnika RCA na rys. 3 przedstawiono fragment rozkładu jego wartości, pomijając ok. 10% obserwacji ekstremalnych. Rozkład wartości wskaźnika RCA (rys. 3) cechuje się znaczną asymetrią prawostronną, natomiast wskaźnik $RCA_k^{(0,2)}$ (rys. 4) ma kształt rozkładu zbliżony do krzywej Gaussa-Laplace'a.



Rys. 4. Rozkład wartości wskaźnika $RCA_k^{(0,2)}$ dla Polski względem Słowacji jako kraju referencyjnego w 2005 r.

Źródło: opracowanie własne na podstawie danych Eurostatu.

W tabeli 1 przedstawiono wyniki testu chi-kwadrat dla rozkładu wartości wskaźnika RCA obliczonego dla Polski względem Słowacji jako kraju referencyjnego (w nawiasach podano wartości prawdopodobieństw testowych).

Na podstawie wyników w tabeli 1 można stwierdzić, że jedynie dwie wartości testu chi-kwadrat są statystycznie istotne i wskazują na stabilność rozkładów empirycznych wskaźnika RCA w następujących okresach: w 2007 r. w porów-

naniu z 2006 r., a także w 2008 r. w porównaniu z 2007 r. W tabeli 2 przedstawiono wartości statystyki chi-kwadrat dla wskaźników $RCA_k^{(1)}$, $RCA_k^{(0,5)}$, $RCA_k^{(0,2)}$ i $RCA_k^{(0,1)}$ obliczonych dla Polski względem Słowacji jako kraju referencyjnego.

Tabela 1. Wyniki testu chi-kwadrat dla rozkładu wartości wskaźnika RCA obliczonego w poszczególnych latach 2005–2009 dla Polski względem Słowacji jako kraju referencyjnego

Lata	2005	2006	2007	2008
2006	33,356 (0,022)	–	–	–
2007	30,867 (0,042)	28,726 (0,070)	–	–
2008	49,646 (0,000)	47,000 (0,000)	20,958 (0,339)	–
2009	41,401 (0,002)	44,310 (0,001)	49,436 (0,000)	37,890 (0,006)

Źródło: opracowanie własne na podstawie danych Eurostatu.

Tabela 2. Wyniki testu chi-kwadrat dla rozkładu wartości wskaźników $RCA_k^{(a)}$ obliczonych w latach 2005–2008 dla Polski względem Słowacji jako kraju referencyjnego

Wskaźnik	$RCA_k^{(1)}$				$RCA_k^{(0,5)}$			
	2005	2006	2007	2008	2005	2006	2007	2008
2006	29,325 (0,061)	–	–	–	21,381 (0,316)	–	–	–
2007	28,759 (0,070)	19,486 (0,426)	–	–	26,912 (0,107)	22,000 (0,284)	–	–
2008	37,060 (0,008)	19,789 (0,407)	28,868 (0,068)	–	30,917 (0,041)	19,409 (0,431)	23,812 (0,203)	–
2009	30,694 (0,044)	17,190 (0,577)	35,668 (0,012)	22,164 (0,276)	20,710 (0,353)	24,660 (0,172)	17,445 (0,560)	18,149 (0,513)
Wskaźnik	$RCA_k^{(0,2)}$				$RCA_k^{(0,1)}$			
	2005	2006	2007	2008	2005	2006	2007	2008
2006	21,381 (0,316)	–	–	–	17,818 (0,535)	–	–	–
2007	16,912 (0,596)	12,959 (0,841)	–	–	17,427 (0,561)	13,813 (0,794)	–	–
2008	12,917 (0,843)	19,409 (0,431)	13,812 (0,795)	–	16,764 (0,606)	16,174 (0,646)	14,843 (0,733)	–
2009	17,710 (0,542)	14,660 (0,744)	22,445 (0,263)	24,149 (0,190)	15,758 (0,673)	15,550 (0,687)	17,038 (0,587)	18,124 (0,514)

Źródło: opracowanie własne na podstawie danych Eurostatu.

Liczba statystycznie istotnych wyników testu chi-kwadrat świadczących o stabilności rozkładu wskaźnika $RCA_k^{(a)}$ w poszczególnych latach wzrasta wraz ze zmniejszaniem wartości parametru a . W przypadku $a = 1$ można stwierdzić 7 wartości statystycznie istotnych (na 10 wyników), dla $a = 0,5$ zaobserwowano 9 wyników statystycznie istotnych, natomiast dla $a = 0,2$ bądź $a = 0,1$ żaden z wyników testu nie był statystycznie istotny.

Z przeprowadzonych obliczeń wynika, że zaproponowany wskaźnik $RCA_k^{(a)}$ cechuje się znacznie wyższym poziomem stabilności w czasie niż klasyczny wskaźnik RCA .

4. Ocena stabilności empirycznego rozkładu wartości wskaźnika przewagi komparatywnej względem grup towarowych

Celem dalszej analizy jest porównanie empirycznych rozkładów wartości wskaźników RCA i $RCA_k^{(a)}$ w wybranych grupach towarowych zgodnych z nomenklaturą międzynarodowej standardowej klasyfikacji handlu (*standard international trade classification* – SITC) oraz ocena wrażliwości wartości charakterystyk liczbowych tych rozkładów względem grup towarowych. W badaniach wzięto pod uwagę grupy towarowe o dostatecznie dużej liczbie obserwacji – wskaźników względnej przewagi komparatywnej obliczonych dla podgrup towarowych wyróżnionych na pięciocyfrowym poziomie dezagregacji. Ostatecznie w badaniach uwzględniono 5 grup towarowych zawierających co najmniej 300 obserwacji:

- SITC 0 – żywność i żywe zwierzęta,
- SITC 5 – chemikalia i produkty pokrewne,
- SITC 6 – artykuły przemysłowe klasyfikowane według materiałów,
- SITC 7 – maszyny, urządzenia i środki transportu,
- SITC 8 – inne wyroby przemysłowe.

Przed wykonaniem analizy ze zbioru obliczonych wskaźników przewagi komparatywnej usunięto obserwacje odstające, które mogły zniekształcać opis rozkładu wskaźnika. W ramach każdej grupy towarowej SITC wyeliminowano 10% wartości ekstremalnych.

W tabelach 3 i 4 przedstawiono wybrane charakterystyki liczbowe w rozkładzie wartości wskaźników RCA i $RCA_k^{(a)}$ dla $a = 0,5$, obliczone dla Polski względem Słowacji jako kraju odniesienia.

Średnie wartości wskaźnika RCA , jak również wartości median przedstawione w tabeli 3 we wszystkich rozważanych grupach towarowych SITC są większe od jedności. Wynika z tego, że Polska posiada przewagę komparatywną w eksporcie względem Słowacji we wszystkich rozważanych grupach towarowych SITC. To spostrzeżenie znajduje również potwierdzenie w wynikach przedstawionych

w tabeli 4 – przeciętne wartości wskaźnika $RCA_k^{(0,5)}$ są większe od 0 we wszystkich grupach towarowych. Wartości współczynników asymetrii zawarte w tabeli 4 pokazują również, iż asymetria w rozkładach wskaźnika $RCA_k^{(0,5)}$ w poszczególnych grupach towarowych SITC jest zazwyczaj słaba i lewostronna, podczas gdy asymetria w rozkładach wskaźnika RCA jest zwykle silna i prawostronna (por. tabela 3).

Tabela 3. Wybrane charakterystyki rozkładu wartości wskaźnika RCA obliczonego według grup towarowych SITC dla Polski względem Słowacji jako kraju referencyjnego w 2009 r.

Charakterystyka liczbowa	SITC 0	SITC 5	SITC 6	SITC 7	SITC 8
Średnia	4,791	4,763	3,754	2,733	2,439
Mediana	2,984	2,252	1,563	1,074	1,415
Minimum	0,443	0,07	0,074	0,079	0,142
Maksimum	24,989	22,698	34,414	21,464	13,081
Kwartyl dolny	1,216	0,69	0,539	0,477	0,583
Kwartyl górny	6,459	6,275	4,297	3,271	3,032
Odchylenie standardowe	5,058	5,746	5,491	3,819	2,707
Współczynnik asymetrii	1,92	1,568	2,73	2,425	1,877

Źródło: opracowanie własne na podstawie danych Eurostatu.

Tabela 4. Wybrane charakterystyki rozkładu wartości wskaźnika $RCA_k^{(0,5)}$ obliczonego według grup towarowych SITC dla Polski względem Słowacji jako kraju referencyjnego w 2009 r.

Charakterystyka liczbowa	SITC 0	SITC 5	SITC 6	SITC 7	SITC 8
Średnia	0,241	0,162	0,097	0,105	0,074
Mediana	0,267	0,2	0,111	0,018	0,086
Minimum	-0,201	-0,581	-0,571	-0,56	-0,452
Maksimum	0,667	0,653	0,709	0,911	0,567
Kwartyl dolny	0,049	-0,093	-0,153	-0,183	-0,134
Kwartyl górny	0,435	0,429	0,349	0,288	0,27
Odchylenie standardowe	0,239	0,331	0,323	0,399	0,264
Współczynnik asymetrii	-0,150	-0,397	-0,157	0,527	-0,095

Źródło: opracowanie własne na podstawie danych Eurostatu.

Aby zbadać, który wskaźnik względnej przewagi komparatywnej generuje rozkłady bardziej wrażliwe na rodzaj grupy towarowej, przeprowadzono testy zgodności rozkładów wskaźników RCA i $RCA_k^{(0,5)}$ dla wszystkich par rozważanych

grup towarowych zgodnych z nomenklaturą SITC. Wyniki testów przedstawiono w tabelach 5 i 6.

Tabela 5. Wyniki testu chi-kwadrat dla rozkładu wartości wskaźnika RCA obliczonego w poszczególnych grupach towarowych SITC dla Polski względem Słowacji jako kraju referencyjnego

Grupa towarowa	SITC 5	SITC 6	SITC 7	SITC 8
SITC 0	50,097 (0,000)	10,235 (0,947)	24,867 (0,165)	86,956 (0,000)
SITC 5	–	51,338 (0,000)	39,277 (0,004)	112,001 (0,000)
SITC 6	–	–	41,866 (0,002)	89,728 (0,000)
SITC 7	–	–	–	113,929 (0,000)

Źródło: opracowanie własne na podstawie danych Eurostatu.

Tabela 6. Wyniki testu chi-kwadrat dla rozkładu wartości wskaźnika $RCA_k^{(0,5)}$ obliczonego w poszczególnych grupach towarowych SITC dla Polski względem Słowacji jako kraju referencyjnego

Grupa towarowa	SITC 5	SITC 6	SITC 7	SITC 8
SITC 0	22,24 (0,272)	14,998 (0,723)	23,101 (0,233)	18,244 (0,506)
SITC 5	–	7,04 (0,994)	15,965 (0,660)	23,1444 (0,231)
SITC 6	–	–	11,54 (0,904)	15,828 (0,669)
SITC 7	–	–	–	56,332 (0,000)

Źródło: opracowanie własne na podstawie danych Eurostatu.

W tabeli 5 przedstawiającej wyniki testu chi-kwadrat dla rozkładów wskaźnika RCA znajdują się 2 wartości statystycznie nieistotne, natomiast w tabeli 6 zawierającej wyniki testu chi-kwadrat dla rozkładu wartości wskaźnika $RCA_k^{(0,5)}$ jest 9 nieistotnych wartości testu. Na podstawie otrzymanych rezultatów badań można przypuszczać, że rozkłady wartości wskaźnika $RCA_k^{(a)}$ w poszczególnych grupach SITC są bardziej stabilne niż w przypadku rozkładów wskaźnika RCA . Można zatem sądzić, że rozkład wartości wskaźnika $RCA_k^{(a)}$ jest mniej wrażliwy na wybór grupy towarowej SITC w porównaniu z wartościami wskaźnika RCA . Dodatkowe badania pokazały, że stabilność rozkładów wartości wskaźników $RCA_k^{(a)}$ jest tym wyższa, im niższy jest poziom parametru a .

5. Podsumowanie

Zaproponowana modyfikacja popularnego wskaźnika ujawnionej względnej przewagi komparatywnej wydaje się do pewnego stopnia eliminować jego wady. Znajduje to potwierdzenie w przedstawionych własnościach wskaźnika $RCA_k^{(a)}$ oraz w wynikach badań empirycznych przeprowadzonych na przykładzie Polski i Słowacji jako kraju referencyjnego. Przede wszystkim wartości skorygowanego wskaźnika RCA są unormowane w przedziale $[-1, 1)$, co umożliwia ich łatwą interpretację. Wartości miernika $RCA_k^{(a)}$ generują rozkłady znacznie bardziej symetryczne niż w przypadku rozkładów wskaźnika RCA . Rozkłady te są również bardziej stabilne w czasie i mniej wrażliwe na rodzaj grupy towarowej. Ułatwia to porównywalność poziomów ujawnionej przewagi komparatywnej w ujęciu czasowym i towarowym. Z kolei za widoczny mankament miernika $RCA_k^{(a)}$ należy uznać jego tendencję do silnego zawężenia zakresu swoich wartości przy bliskim zeru poziomie parametru a . Niemniej przez optymalny dobór wartości a zależny od zakresu zmienności oryginalnego wskaźnika RCA można uzyskać pożądane tempo zbieżności wskaźnika $RCA_k^{(a)}$ do wartości granicznych przedziału $[-1, 1)$.

Literatura

- Balassa B. [1965], *Trade Liberalisation and 'Revealed' Comparative Advantage*, „Manchester School of Economics and Social Studies”, vol. 33.
- Dalum B., Laursen K., Villumsen G. [1998], *Structural Change in OECD Export Specialization Patterns: Despecialization and 'Stickiness'*, „International Review of Applied Economics”, vol. 12(3).
- Hoen A.R., Oosterhaven J. [2006], *On the Measurement of Comparative Advantage*, „The Annals of Regional Science”, vol. 40(3).
- Iapadre P.L. [2001], *Measuring International Specialization*, „International Advances in Economic Research”, vol. 7(2).
- Misala J. [2011], *Międzynarodowa konkurencyjność gospodarki narodowej*, PWE, Warszawa.
- Ricardo D. [1957], *Zasady ekonomii politycznej i opodatkowania*, PWE, Warszawa.
- Sobczyk M. [2000], *Statystyka. Podstawy teoretyczne, przykłady, zadania*, Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, Lublin.
- Yu R., Cai J., Leung P. [2000], *The Normalized Revealed Comparative Advantage Index*, „The Annals of Regional Science”, vol. 43(1).

A Proposal to Modify the Revealed Comparative Index with the Application of Rational Functions

The Revealed Comparative Advantage Index (RCA) is used to estimate the export attractiveness and comparative advantage in the country by industry sectors or commodity groups. This measure has many disadvantages because it ranges from 0 to ∞ .

Its empirical distribution is strongly asymmetric and is not stable in time. The Revealed Comparative Advantage Index distribution is strongly dependent on the number of reference countries, industries and commodity groups taken under consideration. In consequence, some RCA distribution parameters are not stable. These problematic properties make its outcomes impossible to compare across time and space.

The paper proposes transforming the RCA measure into a new index, ranging from -1 to 1 , by using a special class of rational functions. The new measure has a symmetric distribution with a stable mean and is independent of the industry sectors or commodity groups. Empirical distributions of the adjusted RCA index are built on the basis of data on Polish export and import commodities by Standard International Trade Classification (SITC).

Keywords: Revealed Comparative Advantage Index, export, stable distribution, chi-square test.

Jan Tatar

Katedra Matematyki
Uniwersytet Ekonomiczny w Krakowie

Modele wskaźnikowe rynku kapitałowego wykorzystujące funkcję regresji wektorów losowych

Streszczenie

Jedną z ważniejszych kategorii modeli rynku kapitałowego stanowią modele wskaźnikowe. Wyrażają one liniową zależność stóp zwrotu z konkretnych (pojedynczych) aktywów od wybranego zestawu czynników. Czynnikami tymi są na ogół stopy zwrotu odpowiednio konstruowanych portfeli; mogą nimi być np. wybrane indeksy giełdowe.

Kluczowe znaczenie w modelach wskaźnikowych mają współczynniki wrażliwości modelowanej stopy zwrotu na zmiany wybranych czynników. Współczynniki te znane są w teorii i praktyce finansów jako tzw. współczynniki β , a jedną z metod ich wyznaczania jest analiza regresji.

We wcześniejszych pracach autor wykazał, że możliwa jest jednoznaczna konstrukcja funkcji regresji dla dwóch wektorów losowych niekoniecznie o tych samych wymiarach. Wynik ten w niniejszym opracowaniu stanowi dogodny punkt wyjścia do uogólnienia postaci modeli wskaźnikowych, na przypadek gdy wektor wybranych stóp zwrotu jest funkcją innego wektora czynników (np. wektora stóp zwrotu z innych aktywów). Uzyskany w ten sposób współczynnik β z oczywistych powodów będzie miał postać macierzową.

Słowa kluczowe: potęga wektora, momenty rozkładu, funkcja regresji, modele wskaźnikowe.

1. Wprowadzenie

Podstawowym pojęciem, na którym opiera się propozycja przedstawiona w niniejszym opracowaniu, jest potęga wektora w przestrzeni Hilberta zaproponowana m.in. w pracach [Tatar 1993 i 1996b]. Przypomnijmy, że jeżeli $(V, +, \cdot)$ jest dowolną przestrzenią Hilberta nad ciałem skalarów R z iloczynem skalarnym $\langle \cdot | \cdot \rangle: V \times V \rightarrow R$ oraz jeżeli $v \in V$ i $k \in N_0 = N \cup \{0\}$ to k -tą potęgę wektora v definiujemy następująco:

Definicja

$$v^0 = 1 \in R$$

oraz

$$v^k = \begin{cases} v^{k-1} \cdot v & \text{dla } k\text{-nieparzystych} \\ \langle v^{k-1} | v \rangle & \text{dla } k\text{-parzystych} \end{cases}$$

Potęgowanie wektorów umożliwiło z kolei zdefiniowanie w naturalny sposób momentów łącznych dowolnego rzędu dowolnego wektora losowego $\xi: \Omega \rightarrow V$. W przypadku momentów zwykłych przyjęto bowiem $m_k = E(\xi^k)$, zaś w przypadku momentów centralnych $\mu_k = E[(\xi - m_1)^k]$. Uzyskane w ten sposób charakterystyki wielowymiarowych rozkładów prawdopodobieństwa istotnie różnią się od znanych w literaturze tzw. momentów mieszanych. Ponadto z własności potęgi wektora losowego wynika m.in. skalarny charakter momentów łącznych rzędu parzystego oraz wektorowy charakter momentów rzędu nieparzystego (por. np. [Tatar 1993, 1996b i 2009]). Własność ta uzasadnia (a niewątpliwie: sprzyja) wykorzystaniu odpowiednich momentów do konstrukcji miar położenia oraz miar rozproszenia rozkładów wielowymiarowych, takich jak np. wariancja, asymetria, kurtoza czy eksces (por. [Budny 2009, Budny i Tatar 2009, Tatar 1996a i 2006]).

Dla dwóch wektorów losowych $\xi: \Omega \rightarrow V$ oraz $\eta: \Omega \rightarrow V$ o wartościach w tej samej przestrzeni Hilberta zdefiniowano z kolei ich moment mieszany rzędu $k + l$ jako $m_{kl} = E(\xi^k \cdot \eta^l)$. Szczególny przypadek momentu mieszanego wektorów ξ i η , tzn. moment m_{11} nazwano ich kowariancją i przyjęto klasyczne oznaczenie $\text{cov}(\xi, \eta)$.

W praktyce najczęściej wykorzystywanym przypadkiem przestrzeni Hilberta jest s -wymiarowa przestrzeń euklidesowa R^s z naturalnym (euklidesowym) iloczynem skalarnym. Dlatego kolejnym krokiem w rozwijaniu proponowanej koncepcji charakteryzacji wielowymiarowych rozkładów prawdopodobieństwa było sformułowanie pytania o własności łącznego rozkładu wektora (ξ, η) , gdzie $\xi: \Omega \rightarrow R^n$ oraz $\eta: \Omega \rightarrow R^m$ również są wektorami losowymi niekoniecznie o tych samych wymiarach, tzn. niekoniecznie $n = m$. Ważnym zagadnieniem okazało się badanie wzajemnej zależności wektorów ξ i η . W tym kontekście rozważano

w szczególności warunkowe wartości oczekiwane jednego z wektorów losowych przy pewnych założeniach czynionych o drugim z nich. Tak sformułowany problem w prosty sposób prowadził do regresji liniowej dwóch wektorów losowych o dowolnych wymiarach (por. [Budny i Tatar 2012, Najman i Tatar 2010, Osiewalski i Tatar 1999]). W pracy [Budny i Tatar 2012] udowodniono następujące dwa twierdzenia.

Twierdzenie 1

Jeżeli $X: \Omega \rightarrow R^n$ oraz $Y: \Omega \rightarrow R^n$ są wektorami losowymi spełniającymi warunki $D^2 X < +\infty$, $D^2 Y < +\infty$ oraz $D^2 X \neq 0$, to wartość oczekiwana $E[(Y - aX - b)^2]$ osiąga minimum dla $a = \frac{\text{cov}(X, Y)}{D^2 X}$ oraz $b = EY - \frac{\text{cov}(X, Y)}{D^2 X} EX$.

Łatwo można zauważyć, że powyższe twierdzenie jest prostym uogólnieniem na przypadek wielowymiarowy znanego w literaturze sposobu dopasowywania metodą najmniejszych kwadratów prostej regresji dwóch jednowymiarowych zmiennych losowych. Podkreślmy jeszcze raz, że powyższe uogólnienie okazało się możliwe dzięki wykorzystaniu przypomnianej wcześniej koncepcji momentów łącznych.

Jeszcze dalej idącego uogólnienia dostarcza twierdzenie 2.

Twierdzenie 2

Niech $X = (X_1, X_2, \dots, X_m): \Omega \rightarrow R^m$ oraz $Y = (Y_1, Y_2, \dots, Y_n): \Omega \rightarrow R^n$ będą wektorami losowymi takimi, że $X_i, Y_j \in L^2(\Omega) = \left\{ f: \Omega \rightarrow \mathfrak{R} \int_{\Omega} f^2 dP < +\infty \right\}$ oraz $\det \Sigma_X \neq 0$, przy czym Σ_X jest macierzą momentów rzędu drugiego (wariancji-kowariancji) wektora X .

Wówczas wartość oczekiwana $E[(Y - A \cdot X - b)^2]$ osiąga minimum dla $A = \text{Kow}(Y, X) \cdot \Sigma_X^{-1}$ oraz $B = EY - A \cdot EX$, gdzie:

$$\text{Kow}(Y, X) = \begin{bmatrix} E[(Y_1 - EY_1)(X_1 - EX_1)] & \dots & E[(Y_1 - EY_1)(X_m - EX_m)] \\ E[(Y_2 - EY_2)(X_1 - EX_1)] & \dots & E[(Y_2 - EY_2)(X_m - EX_m)] \\ \dots & \dots & \dots \\ E[(Y_n - EY_n)(X_1 - EX_1)] & \dots & E[(Y_n - EY_n)(X_m - EX_m)] \end{bmatrix}.$$

2. Propozycja funkcji regresji finansowych wektorów losowych

Niech $r = (r_1, r_2, \dots, r_m)^T: \Omega \rightarrow R^m$ będzie wektorową zmienną ryzyka (dokładniej: wektorem m zmiennych ryzyka), np. wektorem losowych stóp zwrotu z aktywów

o numerach: $1, 2, \dots, m$. W szczególności możemy przyjąć, że są to rynkowe losowe stopy zwrotu możliwe do osiągnięcia na akcjach spółek: s_1, s_2, \dots, s_m .

Niech z kolei $f = (f_1, f_2, \dots, f_n)^T: \Omega \rightarrow R^n$ będzie n -wymiarowym wektorem czynników ryzyka. Każdy z czynników f_i , dla $i = 1, 2, \dots, n$, jest zatem jedno-wymiarową zmienną losową i może być interpretowany np. jako stopa zwrotu z wybranego indeksu rynkowego bądź jako inny czynnik (na ogół makroekonomiczny), którego wartości – zdaniem modelującego rynek – wpływają na realizacje wektora r , a więc także na wartości jego składowych r_i , gdzie $i = 1, 2, \dots, m$. Zakładamy ponadto, że macierz wariancji-kowariancji wektora f jest nieosobliwa, tzn. $\det \Sigma_f \neq 0$. Założenie to oznacza, że wariancja rozkładu wektora f w sensie Wilksa (za: [Morrison 1990]) jest niezerowa, czyli że rozkład n -wymiarowy nie jest w istocie rozkładem w podprzestrzeni o wymiarze mniejszym niż n . Jeszcze inaczej: nieosobliwość macierzy Σ_f oznacza liniową niezależność współrzędnych (składowych) wektora f .

Będziemy poszukiwać modelu liniowej regresyjnej zależności wektorów r i f , czyli modelu postaci:

$$r = B \cdot f + A + U, \quad (1)$$

gdzie: $A = [\alpha_i]_{m \times 1}$ i $B = [\beta_{ij}]_{m \times n}$ są macierzami rzeczywistymi odpowiedniego wymiaru; zaś $U = [u_i]_{m \times 1}$ jest m -wymiarowym wektorem losowym o zerowym wektorze wartości oczekiwanych spełniającym ponadto warunki:

$$\text{cov}(f_i, u_j) = 0 \text{ dla } i = 1, 2, \dots, n; j = 1, 2, \dots, m \text{ takich, że } i \neq j \quad (2)$$

oraz

$$\text{cov}(u_i, u_j) = 0 \text{ dla } i, j \in \{1, 2, \dots, m\} \text{ takich, że } i \neq j. \quad (3)$$

O ile f jest wektorem wspólnych czynników ryzyka dla wszystkich zmiennych ryzyka, to współrzędne wektora losowego U mogą być interpretowane jako tzw. czynniki specyficzne dla poszczególnych zmiennych ryzyka. Świadczą o tym powyższe założenia o braku skorelowania między różnymi czynnikami specyficznymi oraz między każdym z tych czynników a różniącymi się wskaźnikami (numerami) zmiennymi ryzyka. Innymi słowy, na każdą zmienną ryzyka może mieć wpływ tylko jeden specyficzny czynnik ryzyka (specyficzny właśnie dla danej zmiennej ryzyka).

Specyfikacja modelu (1) polega na znalezieniu postaci macierzy A oraz B . W tym celu wykorzystamy twierdzenie 2. Bezpośrednio z jego tezy wynika, że:

$$B = \text{Kow}(r, f) \cdot \Sigma_f^{-1}, \quad (4)$$

gdzie Σ_f^{-1} jest macierzą (stopnia n) odwrotną do macierzy Σ_f , zaś macierz $\text{Kow}(r, f)$ jest postaci:

$$Kow(r, f) = \begin{bmatrix} E[(r_1 - Er_1)(f_1 - Ef_1)] & \dots & E[(r_1 - Er_1)(f_n - Ef_n)] \\ E[(r_2 - Er_2)(f_1 - Ef_1)] & \dots & E[(r_2 - Er_2)(f_n - Ef_n)] \\ \dots & \dots & \dots \\ E[(r_m - Er_m)(f_1 - Ef_1)] & \dots & E[(r_m - Er_m)(f_n - Ef_n)] \end{bmatrix}_{m \times n}$$

Wymiary macierzy $Kow(r, f)$ oraz \sum_f zapewniają wykonalność mnożenia $Kow(r, f) \cdot \sum_f^{-1}$ i decydują o tym, że jego wynik jest macierzą stopnia $m \times n$, czyli $B = B_{m \times n} = [\beta_{ij}]_{m \times n}$.

Z tezy twierdzenia 2 otrzymujemy również postać macierzy A . Jest bowiem:

$$A = Er - B \cdot Ef. \tag{5}$$

Ponieważ Er (jako wektor wartości oczekiwanych współrzędnych wektora r) jest macierzą wymiaru $m \times 1$, B – macierzą wymiaru $m \times n$, oraz Er – macierzą wymiaru $n \times 1$, więc ostatecznie $A = A_{m \times 1} = [\alpha_{i1}]_{m \times 1}$, a to oznacza, że A jest wektorem m -wymiarowym.

Wobec uzyskanych w (4) oraz (5) postaci macierzy B i A poszukiwany model (równość (1)) przyjmuje postać:

$$r = Kow(r, f) \cdot \sum_f^{-1} \cdot f + Er - Kow(r, f) \cdot \sum_f^{-1} \cdot Ef + U. \tag{6}$$

Powyższa zależność jest równaniem macierzowym, więc w istocie otrzymaliśmy n równań takich, że dla każdego $i = 1, 2, \dots, n$ jest:

$$r_i = \beta_{i1} \cdot f_1 + \beta_{i2} \cdot f_2 + \dots + \beta_{im} \cdot f_m + \alpha_i + u_i, \tag{6'}$$

przy czym

$$\alpha_i = Er_i - \beta_{i1} \cdot Ef_1 - \beta_{i2} \cdot Ef_2 - \dots - \beta_{im} \cdot Ef_m,$$

β_{ij} jest j -tym elementem i -tego wiersza macierzy B (dla każdego $j = 1, 2, \dots, m$), zaś u_j jest i -tym czynnikiem specyficznym spełniającym założenia (2) oraz (3).

3. Przykład funkcji regresji wektorów losowych

Dotychczasowe ogólne rozważania zilustrujemy przykładem, który – przy założeniu niewielkich wymiarów wektorów r oraz f pozwoli lepiej zrozumieć zagregowane macierzowe związki, a ponadto – w rozważanym przypadku – wskaże na efektywną postać współczynników α i β .

Niech zatem $r = (r_1, r_2, r_3): \Omega \rightarrow R^3$ oraz $f = (f_1, f_2, f_3): \Omega \rightarrow R^2$ będą odpowiednio wektorami zmiennych i czynników ryzyka.

O wektorze czynników ryzyka zakładamy ponadto, że macierz jego momentów rzędu drugiego \sum_f jest macierzą nieosobliwą, czyli:

$$\det \sum_f = \det \begin{bmatrix} \text{Var } f_1 & \text{cov}(f_1, f_2) \\ \text{cov}(f_2, f_1) & \text{Var } f_2 \end{bmatrix} = \text{Var } f_1 \cdot \text{Var } f_2 - [\text{cov}(f_1, f_2)]^2 \neq 0.$$

Przy powyższych założeniach poszukujemy regresyjnego liniowego modelu

$$r = B \cdot f + A + U, \quad (1')$$

gdzie:

$$B = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \end{bmatrix}, \quad A = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}, \quad U = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix},$$

przy czym: A i B są macierzami rzeczywistymi, a U jest wektorem losowym, którego współrzędne spełniają warunki: $E(U) = 0$, $\text{cov}(f_i, u_j) = 0$ oraz $\text{cov}(u_i, u_j) = 0$, dla $i \neq j$.

Wobec nieosobliwości macierzy \sum_f istnieje macierz do niej odwrotna o postaci:

$$\sum_f^{-1} = \begin{bmatrix} \frac{\text{Var } f_2}{\text{Var } f_1 \cdot \text{Var } f_2 - [\text{cov}(f_1, f_2)]^2} & \frac{-\text{cov}(f_1, f_2)}{\text{Var } f_1 \cdot \text{Var } f_2 - [\text{cov}(f_1, f_2)]^2} \\ \frac{-\text{cov}(f_2, f_1)}{\text{Var } f_1 \cdot \text{Var } f_2 - [\text{cov}(f_1, f_2)]^2} & \frac{\text{Var } f_1}{\text{Var } f_1 \cdot \text{Var } f_2 - [\text{cov}(f_1, f_2)]^2} \end{bmatrix}.$$

Ponadto:

$$\text{Kow}(r, f) = \begin{bmatrix} \text{cov}(r_1, f_1) & \text{cov}(r_1, f_2) \\ \text{cov}(r_2, f_1) & \text{cov}(r_2, f_2) \\ \text{cov}(r_3, f_1) & \text{cov}(r_3, f_2) \end{bmatrix}.$$

Stąd na mocy tezy twierdzenia 2 mamy:

$$B = \text{Kow}(r, f) \cdot \sum_f^{-1} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \end{bmatrix} = \begin{bmatrix} \frac{\text{cov}(r_1, f_1) \cdot \text{Var } f_2 - \text{cov}(r_1, f_2) \cdot \text{cov}(f_1, f_2)}{\text{Var } f_1 \cdot \text{Var } f_2 - [\text{cov}(f_1, f_2)]^2} & \frac{\text{cov}(r_1, f_2) \cdot \text{Var } f_1 - \text{cov}(r_1, f_1) \cdot \text{cov}(f_1, f_2)}{\text{Var } f_1 \cdot \text{Var } f_2 - [\text{cov}(f_1, f_2)]^2} \\ \frac{\text{cov}(r_2, f_1) \cdot \text{Var } f_2 - \text{cov}(r_2, f_2) \cdot \text{cov}(f_1, f_2)}{\text{Var } f_1 \cdot \text{Var } f_2 - [\text{cov}(f_1, f_2)]^2} & \frac{\text{cov}(r_2, f_2) \cdot \text{Var } f_1 - \text{cov}(r_2, f_1) \cdot \text{cov}(f_1, f_2)}{\text{Var } f_1 \cdot \text{Var } f_2 - [\text{cov}(f_1, f_2)]^2} \\ \frac{\text{cov}(r_3, f_1) \cdot \text{Var } f_2 - \text{cov}(r_3, f_2) \cdot \text{cov}(f_1, f_2)}{\text{Var } f_1 \cdot \text{Var } f_2 - [\text{cov}(f_1, f_2)]^2} & \frac{\text{cov}(r_3, f_2) \cdot \text{Var } f_1 - \text{cov}(r_3, f_1) \cdot \text{cov}(f_1, f_2)}{\text{Var } f_1 \cdot \text{Var } f_2 - [\text{cov}(f_1, f_2)]^2} \end{bmatrix} \quad (5')$$

oraz

$$A = Er - B \cdot Ef = \begin{bmatrix} Er_1 - \beta_{11} \cdot Ef_1 - \beta_{12} \cdot Ef_2 \\ Er_2 - \beta_{21} \cdot Ef_1 - \beta_{22} \cdot Ef_2 \\ Er_3 - \beta_{31} \cdot Ef_1 - \beta_{32} \cdot Ef_2 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}.$$

Ostatecznie poszukiwany model (1') przyjmuje postać:

$$r = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} \beta_{11} \cdot f_1 + \beta_{12} \cdot f_2 + \alpha_1 + u_1 \\ \beta_{21} \cdot f_1 + \beta_{22} \cdot f_2 + \alpha_2 + u_2 \\ \beta_{31} \cdot f_1 + \beta_{32} \cdot f_2 + \alpha_3 + u_3 \end{bmatrix}, \quad (6'')$$

gdzie dla $i = 1, 2, 3$:

$$\beta_{i1} = \frac{\text{cov}(r_i, f_1) \cdot \text{Var} f_2 - \text{cov}(r_i, f_2) \cdot \text{cov}(f_1, f_2)}{\text{Var} f_1 \cdot \text{Var} f_2 - [\text{cov}(f_1, f_2)]^2},$$

$$\beta_{i2} = \frac{\text{cov}(r_i, f_2) \cdot \text{Var} f_1 - \text{cov}(r_i, f_1) \cdot \text{cov}(f_1, f_2)}{\text{Var} f_1 \cdot \text{Var} f_2 - [\text{cov}(f_1, f_2)]^2},$$

$$\alpha_i = Er_i - \frac{\text{cov}(r_i, f_1) \cdot \text{Var} f_2 - \text{cov}(r_i, f_2) \cdot \text{cov}(f_1, f_2)}{\text{Var} f_1 \cdot \text{Var} f_2 - [\text{cov}(f_1, f_2)]^2} \cdot Ef_1 -$$

$$- \frac{\text{cov}(r_i, f_2) \cdot \text{Var} f_1 - \text{cov}(r_i, f_1) \cdot \text{cov}(f_1, f_2)}{\text{Var} f_1 \cdot \text{Var} f_2 - [\text{cov}(f_1, f_2)]^2} \cdot Ef_2.$$

4. Uwagi końcowe

Uzyskany w równaniu (6) – w przypadku szczególnym w równaniu (6'') – model jest, podobnie jak znane w literaturze finansowej modele wielowskaźnikowe (por. [Ross 1976, Sharpe 1963]), dogodnym narzędziem do badania zależności wybranych zmiennych ryzyka od ustalonego uprzednio zestawu czynników ryzyka. Tym razem możliwe staje się badanie tej zależności jednocześnie dla wszystkich zmiennych, a macierz (5) – w przykładzie: (5') – dostarcza stosunkowo prostego sposobu wyznaczania współczynników β . Ich interpretacja również jest prosta i naturalna: β_{ij} jest miarą wrażliwości i -tej zmiennej ryzyka (czyli r_i) na zmiany wartości j -tego czynnika ryzyka (czyli f_j), β_{ij} jest bowiem odpowiednią pochodną cząstkową funkcji $r_i = r_i(f_1, f_2, \dots, f_n)$, tzn. $\beta_{ij} = \frac{\partial r_i}{\partial f_j}$.

Podkreślmy jeszcze raz, że komentowany wynik uzyskano, wykorzystując nową koncepcję tzw. momentów łącznych wielowymiarowych rozkładów prawdopodobieństwa.

Zauważmy ponadto, że w przypadku szczególnym, gdy będziemy uwzględniać tylko jeden czynnik ryzyka f_1 , to każde z równań (6') przyjmie postać $r_i = \beta_{i1} \cdot f_1 + \alpha_i + u_i$ (dla $i = 1, 2, \dots, m$), czyli okaże się tożsamy z zaproponowanym przez W.F. Sharpe'a [1963] jednowskaźnikowym modelem rynku kapitałowego.

Warto na koniec zapytać o postać uzyskanego modelu, w przypadku gdy – jak to czynią niektórzy autorzy (por. np. [Rynki... 2008]) – założymy brak skorelowania między każdymi dwoma czynnikami ryzyka, czyli że $\text{cov}(f_i, f_j) = 0$, gdy $i \neq j$. Wówczas, co szczególnie wyraźnie widać w macierzy (5') przy założeniu $\text{cov}(f_1, f_2) = 0$, współczynniki wrażliwości są postaci $\beta_{ij} = \frac{\text{cov}(r_i, f_j)}{\text{Var} f_j}$.

Literatura

- Budny K. [2009], *Kurtoza wektora losowego*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, nr 78, seria: Ekonometria, nr 26.
- Budny K., Tatar J. [2009], *Kurtosis of a Random Vector – Special Types of Distributions*, „Statistics in Transition – New Series”, vol. 10, nr 3.
- Budny K., Tatar J. [2012], *Regresja liniowa z wykorzystaniem nowej definicji momentów wektorów losowych*, „Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie”, nr 892.
- Morrison D.F. [1990], *Wielowymiarowa analiza statystyczna*, PWN, Warszawa.
- Najman P., Tatar J. [2010], *Regresja wektorów losowych dla wielowymiarowego rozkładu normalnego* [w:] *Badania ekonometryczne w teorii i praktyce*, red. A.S. Barczak, Katowice.
- Osiewalski J., Tatar J. [1999], *Multivariate Chebyshev Inequality Based on a New Definition of Moments of a Random Vector*, „Przegląd Statystyczny”, nr 2.
- Ross S.A. [1976], *The Arbitrage Theory of Capital Asset Pricing*, „Journal of Economic Theory”, vol. 59.
- Rynki, instrumenty i instytucje finansowe* [2008], red. J. Czekaj, PWN, Warszawa.
- Sharpe W.F. [1963], *A Simplified Model for Port Folio Analysis*, „Management Science”, vol. 19.
- Tatar J. [1993], *Moments of a Random Variable in a Hilbert Space*, „Discussion Paper”, nr 1, Cracow Academy of Economics (także w: „Przegląd Statystyczny” 1999, nr 2).
- Tatar J. [1996a], *Nierówność Czebyszewa dla wielowymiarowych zmiennych losowych*, „Badania Operacyjne i Decyzje”, nr 2.
- Tatar J. [1996b], *O niektórych miarach rozproszenia rozkładów prawdopodobieństwa*, „Przegląd Statystyczny”, z. 3–4.
- Tatar J. [2006], *Późnieźmienniki i momenty w charakteryzacji wielowymiarowych rozkładów prawdopodobieństwa* [w:] *Matematyka – język uniwersalny*, Księga jubileuszowa dla uczczenia 70. urodzin Profesora Tadeusza Stanisza, red. E. Smaga, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.
- Tatar J. [2009], *Nowe charakterystyki warunkowych rozkładów wielowymiarowych*, Studia i Prace Uniwersytetu Ekonomicznego w Krakowie, nr 3.

Indicator Models of the Capital Market that Use the Regression Function of Random Vectors

One of the more important categories of the capital market, indicator models show the linear dependence of rate from specific (individual) assets on a selected set of factors. These factors are usually the profitability of appropriately constructed portfolios; these can include, for example, specific stock indexes. The coefficients of the sensitivity of the rates on the changes of specific factors are essential for the indicator models. In financial theory and practice those coefficients are known as beta coefficients and one method of determining them is regression analysis.

In the previous works the author showed that construction of the regression function of two random vectors – not necessarily of the same dimensions – is possible. That result – in this paper – is a convenient starting point to generalising the indicator model form for cases where the vector of selected repayment rates is a function of other vector factors (e.g. the repayment rate vector of other assets). The beta coefficient obtained in this way, for obvious reasons, will have a matrix form.

Keywords: power of a vector, moments of a distribution, regression function, indicator models.

Katarzyna Budny

Katedra Matematyki

Uniwersytet Ekonomiczny w Krakowie

Wybrane własności kurtozy wektora losowego

Streszczenie

Prezentowana praca jest kontynuacją rozważań prowadzonych nad kurtozą wektora losowego. Kurtoza wektora losowego rozumiana jest tutaj jako moment centralny czwartego rzędu wektora losowego podzielony przez kwadrat jego wariancji. Pojęcie momentu centralnego wektora losowego, a w szczególności wariancji, opiera się na definicji potęgi wektora i zostało zaproponowane przez J. Tatara. W niniejszym artykule przedstawione zostały wybrane, istotne własności kurtozy wektora losowego. Należą do nich m.in. własność niezmienniczości względem pewnych przekształceń afinicznych. Ponadto ustalony został związek między kurtozą wektora losowego a kwadratem jego współczynnika asymetrii. Spełnienie podanych własności przez tak skonstruowaną miarę może uzasadniać jej wybór na wielowymiarowy odpowiednik kurtozy jednowymiarowej zmiennej losowej.

Słowa kluczowe: kurtoza wektora losowego, potęga wektora, rozkład wielowymiarowy, charakterystyki rozkładu wielowymiarowego.

1. Wprowadzenie

Prezentowana praca jest kontynuacją rozważań prowadzonych nad kurtożą wektora losowego [Budny i Tatar 2009, Budny 2009]. U podstaw tych rozważań leży zaproponowana przez J. Tatara [1996 i 1999] definicja potęgi wektora w przestrzeni z iloczynem skalarnym.

Definicja 1. Dla dowolnego $v \in R^n$ oraz dowolnej liczby $k \in N_0 = N \cup \{0\}$ k -tą potęgę wektora v definiujemy w następujący sposób:

$$v^o = 1 \in R \text{ oraz } v^k = \begin{cases} v^{k-1} \cdot v & \text{dla } k\text{-nieparzystych} \\ \langle v^{k-1}, v \rangle & \text{dla } k\text{-parzystych} \end{cases}$$

W pracy rozważać będziemy przestrzeń wektorową $(R^n, R, +, \cdot)$ w której określono klasyczny (euklidesowy) iloczyn skalarny postaci:

$$\langle v, w \rangle := \sum_{i=1}^n v_i w_i, \text{ gdzie } v = (v_1, v_2, \dots, v_n), w = (w_1, w_2, \dots, w_n) \in R^n.$$

Dzięki wykorzystaniu pojęcia potęgi wektora zostały zdefiniowane m.in. momenty centralne wielowymiarowego wektora losowego [Tatar 1996 i 1999].

W pracy przyjęto następujące oznaczenie:

$-D^2X$ to moment centralny rzędu 2 wektora losowego $X = (X_1, \dots, X_k): \Omega \rightarrow R^k$ oparty na definicji potęgi wektora czy też równoważnie wariancja całkowita, czyli $D^2X = \sum_{i=1}^k D^2X_i$ (por. [Bilodeu i Brenner 1999]).

Za pomocą momentów centralnych wektora losowego, na podstawie postaci kurtozy jednowymiarowej zmiennej losowej [Cramer 1958, Jakubowski i Sztencel 2004], zdefiniowana została kurtoza wektora losowego.

Założmy zatem, że $X = (X_1, \dots, X_n): \Omega \rightarrow R^n$ jest wektorem losowym, dla którego istnieje moment centralny czwartego rzędu.

Definicja 2. Przez kurtozę wektora losowego $X: \Omega \rightarrow R^n$ rozumiemy liczbę postaci:

$$\beta_{2,n} = \text{Kurt}X = \frac{E[(X - EX)^4]}{(D^2X)^2}.$$

Tak skonstruowany wskaźnik posiada szereg istotnych, pożądaných własności. Dotychczas podano postać kurtozy dla wybranych typów rozkładów (przypadek wektorów losowych o stochastycznie niezależnych współrzędnych, posiadających rozkłady normalne czy też rozkłady t -Studenta) [Budny i Tatar 2009] oraz nieco ogólniej wyznaczono postać kurtozy dla wektora losowego o stochastycznie niezależnych współrzędnych [Budny 2009]. Warto w tym miejscu zwrócić uwagę na wektor losowy, którego wszystkie współrzędne posiadają ten sam rozkład normalny i są stochastycznie niezależne. Kurtoza takiego wektora losowego jest bowiem postaci:

$$\beta_{2,n} = 1 + \frac{2}{n}.$$

W pracy omówione zostaną kolejne ważne własności kurtozy wektora losowego, rozumianej jako iloraz momentu centralnego czwartego rzędu wektora losowego dzielonego przez kwadrat jego wariancji całkowitej. Własności te wraz z przedstawionymi powyżej uzasadniają wybór tak skonstruowanej miary na wielowymiarowy odpowiednik kurtozy zmiennej losowej (jednowymiarowej).

W literaturze przedmiotu pojawiają się propozycje definicji kurtozy wektora losowego [Mardia 1970, Kotz, Johnston i Balakrishnan 2000]. Prezentowane w niniejszym opracowaniu podejście jest inne niż spotykane do tej pory.

2. Kurtoza jako niezmiennik pewnych przekształceń afinicznych

W kolejnych trzech twierdzeniach zaprezentowana zostanie własność niezmienniczości kurtozy względem wybranych typów przekształceń afinicznych.

Twierdzenie 1. Kurtoza jako niezmiennik skali

Przypuśćmy, że $X = (X_1, \dots, X_n): \Omega \rightarrow \mathfrak{R}^n$ jest wektorem losowym, dla którego istnieje kurtoza, natomiast $a \in \mathfrak{R}$ jest dowolną liczbą rzeczywistą. Wtedy:

$$\text{Kurt}(aX) = \text{Kurt}X. \quad (1)$$

Dowód. Poniższe przekształcenia uzasadniające równość (1) wynikają z definicji wariancji wektora losowego, definicji parzystych potęg wektora losowego oraz własności iloczynu skalarnego:

$$\begin{aligned} \text{Kurt}(aX) &= \frac{E[(aX - E(aX))^4]}{(D^2(aX))^2} = \frac{E(\langle aX - E(aX), aX - E(aX) \rangle^2)}{(E(\langle aX - E(aX), aX - E(aX) \rangle))^2} = \\ &= \frac{E(a^4 \langle X - EX, X - EX \rangle^2)}{(a^2 (D^2 X))^2} = \frac{a^4 E(\langle X - EX, X - EX \rangle^2)}{a^4 (D^2 X)^2} = \frac{E[(X - EX)^4]}{(D^2 X)^2} = \text{Kurt}X. \end{aligned}$$

Twierdzenie 2. Kurtoza jako niezmiennik translacji

Niech $X: \Omega \rightarrow R^n$ będzie wektorem losowym, dla którego istnieje kurtoza, oraz niech $b \in R^n$ będzie dowolnym wektorem z przestrzeni R^n . Wówczas:

$$\text{Kurt}(X + b) = \text{Kurt}X. \quad (2)$$

Dowód. Równość (2) wynika wprost z definicji kurtozy wektora losowego oraz z faktu, że wariancja wektora losowego jest niezmiennikiem translacji [Tatar 1999]. Istotnie:

$$\begin{aligned} \text{Kurt}(X + b) &= \frac{E[(X + b - E(X + b))^4]}{(D^2(X + b))^2} = \frac{E[(X + b - E(X) - b)^4]}{(D^2 X)^2} = \\ &= \frac{E[(X - E(X) - b)^4]}{(D^2 X)^2} = \frac{E[(X - E(X))^4]}{(D^2 X)^2} = \text{Kurt}X. \end{aligned}$$

Przed sformułowaniem kolejnej własności kurtozy wektora losowego rozpatrzymy następujący lemat.

Lemat 1. Załóżmy, że $X: \Omega \rightarrow R^n$ jest wektorem losowym, dla którego istnieje moment centralny rzędu $2k$ oraz $C_{(n \times n)}$ jest macierzą ortogonalną (tzn. $C^T = C^{-1}$).

Wówczas:

$$E[(CX - E(CX))^{2k}] = E[(X - E(X))^{2k}],$$

czyli momenty centralne parzystego rzędu są niezmiennikami przekształceń ortogonalnych.

Dowód. Na początek zauważmy, że macierz CX jest postaci:

$$CX = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n c_{1j} X_j \\ \vdots \\ \sum_{j=1}^n c_{nj} X_j \end{bmatrix}.$$

Dzięki własnościom momentów centralnych parzystego rzędu otrzymujemy:

$$\begin{aligned} E[(CX - E(CX))^{2k}] &= E\left[\left((C(X - EX))^{2k}\right)\right] = E\left[\left(\sum_{i=1}^n \left(\sum_{j=1}^n c_{ij} (X_j - E(X_j))\right)^2\right)^k\right] = \\ &= E\left[\sum_{i_1, \dots, i_k=1}^n \left(\sum_{s, t=1}^n c_{i_1 s} c_{i_1 t} (X_s - E(X_s))(X_t - E(X_t))\right) \dots \left(\sum_{s, t=1}^n c_{i_k s} c_{i_k t} (X_s - E(X_s))(X_t - E(X_t))\right)\right] = \\ &= E\left[\sum_{i_1, \dots, i_k=1}^n \left(\sum_{s_1, \dots, s_k, t_1, \dots, t_k=1}^n c_{i_1 s_1} c_{i_1 t_1} \dots c_{i_k s_k} c_{i_k t_k} (X_{s_1} - E(X_{s_1}))(X_{t_1} - E(X_{t_1})) \dots \right.\right. \\ &\quad \left.\left. \dots (X_{s_k} - E(X_{s_k}))(X_{t_k} - E(X_{t_k}))\right)\right] = \\ &= E\left[\sum_{s_1, \dots, s_k, t_1, \dots, t_k=1}^n (X_{s_1} - E(X_{s_1}))(X_{t_1} - E(X_{t_1})) \dots (X_{s_k} - E(X_{s_k}))(X_{t_k} - E(X_{t_k})) \cdot \right. \\ &\quad \left. \cdot \left(\sum_{i_1, \dots, i_k=1}^n c_{i_1 s_1} c_{i_1 t_1} \dots c_{i_k s_k} c_{i_k t_k}\right)\right] = \\ &= E\left[\sum_{s_1, \dots, s_k, t_1, \dots, t_k=1}^n (X_{s_1} - E(X_{s_1}))(X_{t_1} - E(X_{t_1})) \dots (X_{s_k} - E(X_{s_k}))(X_{t_k} - E(X_{t_k})) \cdot \right. \\ &\quad \left. \cdot \left(\sum_{i_1=1}^n c_{i_1 s_1} c_{i_1 t_1}\right) \dots \left(\sum_{i_k=1}^n c_{i_k s_k} c_{i_k t_k}\right)\right] = \end{aligned}$$

$$= E \left[\sum_{s[1], \dots, s[k], t[1], \dots, t[k]=1}^n \left(X_{s[1]} - E(X_{s[1]}) \right) \left(X_{t[1]} - E(X_{t[1]}) \right) \cdots \left(X_{s[k]} - E(X_{s[k]}) \right) \left(X_{t[k]} - E(X_{t[k]}) \right) \cdot \langle c_{\cdot, s[1]}, c_{\cdot, t[1]} \rangle \cdots \langle c_{\cdot, s[k]}, c_{\cdot, t[k]} \rangle \right],$$

gdzie $c_{\cdot j}$ to j -ta kolumna macierzy CX .

Przypomnijmy, że kolumny (wiersze) macierzy ortogonalnej tworzą układ ortonormalny, a zatem:

$$E[(CX - E(CX))^{2k}] = E \left[\sum_{s[1], \dots, s[k], t[1], \dots, t[k]=1}^n \left(X_{s[1]} - E(X_{s[1]}) \right) \left(X_{t[1]} - E(X_{t[1]}) \right) \cdots \left(X_{s[k]} - E(X_{s[k]}) \right) \left(X_{t[k]} - E(X_{t[k]}) \right) \cdot I_{s[1]t[1]} \cdots I_{s[k]t[k]} \right], \text{ gdzie } I_{st} = \begin{cases} 1 & \text{dla } s = t \\ 0 & \text{dla } s \neq t \end{cases}$$

Wobec tego:

$$\begin{aligned} E[(CX - E(CX))^{2k}] &= E \left[\sum_{s[1], \dots, s[k]=1}^n \left(X_{s[1]} - E(X_{s[1]}) \right)^2 \cdots \left(X_{s[k]} - E(X_{s[k]}) \right)^2 \right] = \\ &= E \left[\left(\sum_{s[1]=1}^n \left(X_{s[1]} - E(X_{s[1]}) \right)^2 \right) \cdots \left(\sum_{s[k]=1}^n \left(X_{s[k]} - E(X_{s[k]}) \right)^2 \right) \right] = E \left[\left(\sum_{s=1}^n (X_s - E(X_s))^2 \right)^k \right] = \\ &= E[(X - E(X))^{2k}]. \end{aligned}$$

Twierdzenie 3. Kurtoza jako niezmiennik przekształceń ortogonalnych

Jeżeli $X: \Omega \rightarrow R^n$ jest wektorem losowym, dla którego istnieje kurtoza, oraz $C_{(n \times n)}$ jest macierzą ortogonalną (tzn. $C^T = C^{-1}$) to:

$$\text{Kurt}(CX) = \text{Kurt}X.$$

Dowód. Twierdzenie 3 jest natychmiastowym wnioskiem z lematu 1. Istotnie z lematu 1 wynika, że momenty centralne czwartego oraz drugiego rzędu (wariancja) są niezmiennikami przekształceń ortogonalnych, a zatem kurtoza wektora losowego jako iloraz ich odpowiednich funkcji również spełnia tę własność.

W następnym rozdziale ustalone zostaną ograniczenia dolne dla wartości kurtozy.

3. Kurtoza wektora losowego – ograniczenia dolne

Twierdzenie 4. Niech $X: \Omega \rightarrow R^n$ będzie wektorem losowym, dla którego istnieje kurtoza $\beta_{2,n}$. Wówczas:

$$\beta_{2,n} \geq 1.$$

Dowód. Kluczową rolę w dowodzie odgrywać będzie nierówność Jensena. Przypomnijmy zatem jej postać:

$$g(E[\tilde{X}]) \leq E[g(\tilde{X})],$$

gdzie \tilde{X} to zmienna losowa, natomiast g to funkcja wypukła.

Zauważmy, że $g(x) = x^2$ jest funkcją wypukłą. Jeżeli w nierówności Jensena za \tilde{X} przyjmiemy zmienną losową $(X - m)^2$, to uzyskamy:

$$(E[(X - m)^2])^2 \leq E[((X - m)^2)^2],$$

czyli:

$$(D^2 X)^2 \leq E[(X - m)^4].$$

Mamy więc $\frac{E[(X - m)^4]}{(D^2 X)^2} \geq 1$, a zatem $\beta_{2,n} \geq 1$.

Uwaga 1. Ograniczenie dolne dla kurtozy wektora losowego w przypadku n -wymiarowym jest realizowane przez rozkład wektora losowego, będącego zestawieniem n niezależnych zmiennych losowych o tym samym rozkładzie dwupunktowym (zero-jedynkowym) z $p = \frac{1}{2}$, czyli przez szczególny przypadek wielowymiarowego rozkładu dwupunktowego.

Dowód. Istotnie dla wektora losowego $X = (X_1, \dots, X_n): \Omega \rightarrow \mathfrak{R}^n$ spełniającego powyższe założenia kurtoza jest postaci:

$$\begin{aligned} \beta_{2,n} &= \frac{\sum_{i=1}^n E[(X_i - m_i)^4] + \sum_{\substack{i,j=1 \\ i \neq j}}^n E[(X_i - m_i)^2]E[(X_j - m_j)^2]}{(D^2 X)^2} = \\ &= \frac{n(D^2 X_i) \text{Kurt} X_i + n(n-1)(D^2 X_i)^2}{n^2(D^2 X_i)^2} = \frac{\text{Kurt} X_i + n-1}{n}, \end{aligned}$$

gdzie X_i to zmienna losowa o rozkładzie zero-jedynkowym z $p = \frac{1}{2}$, dla której kurtoza wynosi $\frac{(3p^2 - 3p + 1)}{p(1-p)}$, czyli $\text{Kurt}X_i = \frac{\left(3\left(\frac{1}{2}\right)^2 - 3\left(\frac{1}{2}\right) + 1\right)}{\frac{1}{2}\left(1 - \frac{1}{2}\right)} = 1$. Wobec tego $\beta_{2,n} = \frac{1+n-1}{n} = 1$.

Zauważmy ponadto, że pośród wektorów losowych będących zestawieniem n niezależnych zmiennych losowych o tym samym rozkładzie dwupunktowym (zero-jedynkowym) jedynie wektor z $p = \frac{1}{2}$, ma kurtozę równą 1.

Istotnie z poprzednich obliczeń wynika, iż:

$$\beta_{2,n} = \frac{\frac{(3p^2 - 3p + 1)}{p(1-p)} + n - 1}{n} = \frac{(3p^2 - 3p + 1) + (n-1)p(1-p)}{np(1-p)}. \quad (3)$$

Szukamy $p \in (0, 1)$ spełniających równanie $\beta_{2,n} = 1$, a więc:

$$\frac{(3p^2 - 3p + 1) + (n-1)p(1-p)}{np(1-p)} = 1.$$

Otrzymujemy wobec tego równanie kwadratowe zmiennej p postaci: $4p^2 - 4p + 1 = 0$, którego jedynym rozwiązaniem jest $p = \frac{1}{2}$.

Uwaga 2. Nie ma górnego ograniczenia dla wartości kurtozy n -wymiarowego wektora losowego.

Dowód. Dla wektorów losowych będących zestawieniem n niezależnych zmiennych losowych o tym samym rozkładzie dwupunktowym (zero-jedynkowym) z postaci (3) wynika, że jeżeli $p \rightarrow 0^+$ lub $p \rightarrow 1^-$, to $\beta_{2,n} \rightarrow +\infty$.

Przed sformułowaniem kolejnego twierdzenia przypomnijmy następującą definicję.

Definicja 3. Współczynnikiem asymetrii wektora losowego $X: \Omega \rightarrow R^n$ nazywamy wektor postaci [Tatar 2000]:

$$\gamma_{1,n} = \text{Skew}X = \frac{E[(X - EX)^3]}{(D^2 X)^{\frac{3}{2}}}.$$

Twierdzenie 5. Nierówność między kurtozą wektora losowego a kwadratem jego współczynnika asymetrii

Niech $X = (X_1, \dots, X_n): \Omega \rightarrow \mathfrak{R}^n$ będzie wektorem losowym, dla którego istnieje kurtoza. Załóżmy, że wartość oczekiwana tego wektora to m , a macierz kowariancji ma postać $\sigma^2 I$, gdzie $\sigma > 0$. Wówczas prawdziwa jest nierówność:

$$\beta_{2,n} \geq 1 + n(\gamma_{1,n})^2. \quad (4)$$

Dowód. W pierwszej części dowodu wykażemy, że nierówność (4) jest prawdziwa dla wektorów losowych (spełniających założenia twierdzenia 5) o zerowym wektorze wartości oczekiwanych i jednostkowej macierzy kowariancji.

Niech zatem $X \sim d_n(0, I)$. Z tego faktu (w oczywisty sposób) wynikają następujące własności:

$$- E(X_i) = 0 \text{ dla każdego } i \in \{1, \dots, n\}, \quad (5)$$

$$- D^2 X_i = 1 \text{ dla każdego } i \in \{1, \dots, n\} \Rightarrow D^2 X = E(X^2) = n, \quad (6)$$

$$- E(X_i X_j) = 0 \text{ dla wszystkich } i, j \in \{1, \dots, n\} \text{ takich, że } i \neq j. \quad (7)$$

Zauważmy, że dla wszystkich $a_0 \in R, a_1 \in R^n$ oraz $a_2 \in R$ zachodzi nierówność:

$$E[(a_0 + \langle a_1, X \rangle + a_2)^2] \geq 0. \quad (8)$$

Nierówność (8) można natomiast przedstawić w postaci:

$$a_0^2 + \langle a_1, a_1 \rangle + 2a_0 a_2 n + 2a_2 E[\langle a_1, X \rangle X^2] + a_2^2 E[X^4] \geq 0. \quad (9)$$

Istotnie z własności całki uzyskujemy:

$$\begin{aligned} E[(a_0 + \langle a_1, X \rangle + a_2 X^2)^2] &= E[(a_0 + \langle a_1, X \rangle)^2 + 2(a_0 + \langle a_1, X \rangle)a_2 X^2 + (a_2 X^2)^2] = \\ &= E[a_0^2 + 2a_0 \langle a_1, X \rangle + \langle a_1, X \rangle^2 + 2a_0 a_2 X^2 + 2\langle a_1, X \rangle a_2 X^2 + a_2^2 X^4] = \\ &= a_0^2 + 2a_0 E[\langle a_1, X \rangle] + E[\langle a_1, X \rangle^2] + 2a_0 a_2 E[X^2] + 2a_2 E[\langle a_1, X \rangle X^2] + a_2^2 E[X^4]. \end{aligned}$$

Zwróćmy teraz uwagę na równości:

$$E[\langle a_1, X \rangle] = 0, \quad (10)$$

$$E[\langle a_1, X \rangle^2] = \langle a_1, a_1 \rangle. \quad (11)$$

Wynikają one bezpośrednio z własności (5)–(7).

$$\text{Rzeczywiście } E[\langle a_1, X \rangle] = E\left(\sum_{i=1}^n a_i^1 X_i\right) = \sum_{i=1}^n a_i^1 E(X_i) \stackrel{(5)}{=} 0,$$

natomiast:

$$\begin{aligned} E[\langle a_1, X \rangle^2] &= E\left[\left(\sum_{i=1}^n a_i^1 X_i\right)^2\right] = E\left[\sum_{i,j=1}^n a_i^1 a_j^1 X_i X_j\right] = \sum_{i=1}^n (a_i^1)^2 E(X_i^2) + \sum_{\substack{i,j=1 \\ i \neq j}}^n a_i^1 a_j^1 E(X_i X_j) \stackrel{(6),(7)}{=} \\ &= \sum_{i=1}^n (a_i^1)^2 = \langle a_1, a_1 \rangle. \end{aligned}$$

Dzięki równościom (10) i (11) otrzymujemy zatem postać (9).
Zauważmy również, że zachodzą następujące równości:

$$(E(X^3))^2 = \sum_{i=1}^n \left(\sum_{j=1}^n E(X_j^2 X_i) \right)^2 \quad (12)$$

oraz

$$E[\langle E(X^3), X \rangle X^2] = (E(X^3))^2. \quad (13)$$

Istotnie:

$$\begin{aligned} (E(X^3))^2 &= \langle E(X^3), E(X^3) \rangle = \\ &= \left\langle \left(E\left(\sum_{j=1}^n X_j^2 X_1 \right), \dots, E\left(\sum_{j=1}^n X_j^2 X_n \right) \right), \left(E\left(\sum_{j=1}^n X_j^2 X_1 \right), \dots, E\left(\sum_{j=1}^n X_j^2 X_n \right) \right) \right\rangle = \\ &= \left\langle \left(\sum_{j=1}^n E(X_j^2 X_1), \dots, \sum_{j=1}^n E(X_j^2 X_n) \right), \left(\sum_{j=1}^n E(X_j^2 X_1), \dots, \sum_{j=1}^n E(X_j^2 X_n) \right) \right\rangle = \sum_{i=1}^n \left(\sum_{j=1}^n E(X_j^2 X_i) \right)^2. \end{aligned}$$

Do równości (13) natomiast prowadzą poniższe przekształcenia, w których korzystamy z własności (12):

$$\begin{aligned} E[\langle E(X^3), X \rangle X^2] &= E \left[\left\langle \left(E\left(\sum_{j=1}^n X_j^2 X_1 \right), \dots, E\left(\sum_{j=1}^n X_j^2 X_n \right) \right), (X_1, \dots, X_n) \right\rangle \left(\sum_{s=1}^n X_s^2 \right) \right] = \\ &= E \left[\sum_{i=1}^n \left(\sum_{j=1}^n E(X_j^2 X_i) X_i \right) \left(\sum_{s=1}^n X_s^2 \right) \right] = E \left[\sum_{i,s=1}^n \left(\sum_{j=1}^n E(X_j^2 X_i) X_i \right) X_s^2 \right] = \sum_{i,s=1}^n \left(\sum_{j=1}^n E(X_j^2 X_i) \right) E(X_i X_s^2) = \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n E(X_j^2 X_i) \right) \left(\sum_{s=1}^n E(X_s^2 X_i) \right) = \sum_{i=1}^n \left(\sum_{j=1}^n E(X_j^2 X_i) \right)^2 \stackrel{(12)}{=} (E(X^3))^2. \end{aligned}$$

Nierówność (9) jest prawdziwa dla wszystkich $a_0 \in R, a_1 \in R^n$, oraz $a_2 \in R$, a więc w szczególności dla a_2 spełniającego równość $a_2^2 = \frac{1}{n^2}$ oraz a_0 takiego, że $a_0^2 + 2a_0 a_2 n = -1$.

Przyjmijmy, że $a_2 = \frac{1}{n}$. Wówczas funkcja $f(a_0) = a_0^2 + 2a_0 \frac{1}{n} n = a_0^2 + 2a_0$ osiąga minimum w punkcie $a_0 = -1$ równe $f_{\min}(-1) = -1$. Przyjmijmy zatem $a_0 = -1$. Zauważmy, że (przy powyższych a_0 i a_2) jeśli $a_1 = cE(X^3)$, gdzie $c \in R$, to na podstawie własności (13) nierówność (9) przybiera postać:

$$-1 + c^2 \langle E(X^3), E(X^3) \rangle + \frac{2c}{n} (E(X^3))^2 + \frac{E[X^4]}{n^2} \geq 0.$$

Mamy zatem:

$$-1 + \left(\frac{c^2 n + 2c}{n} \right) (E(X^3))^2 + \beta_{2,n} \geq 0.$$

Przy założeniu $X \sim d_n(0, I)$ otrzymujemy $(\gamma_{1,n})^2 = \frac{(E(X^3))^2}{n^3}$. Wówczas

$$-1 + (c^2 n^3 + 2cn^2)(\gamma_{1,n})^2 + \beta_{2,n} \geq 0.$$

Stałą $c \in R$ chcemy tak dobrać, aby $c^2 n^3 + 2cn^2 = -n$. Otrzymujemy więc równanie kwadratowe zmiennej c postaci $n^3 c^2 + 2n^2 c + n = 0$, w którym $\Delta = 0$, czyli posiadające (w zbiorze liczb rzeczywistych) jedyne rozwiązanie postaci $c_0 = \frac{-2n^2}{2n^3} = -\frac{1}{n}$. Za a_1 przyjmujemy zatem wektor postaci $-\frac{1}{n}E(X^3)$.

Podsumowując, dla $a_0 = -1$, $a_1 = -\frac{1}{n}E(X^3)$ oraz $a_2 = \frac{1}{n}$ nierówność (9) przyjmuje postać nierówności (4) (przy założeniu $X \sim d_n(0, I)$).

Weźmy teraz pod uwagę wektor losowy X o wektorze wartości oczekiwanych m oraz macierzy kowariancji postaci $\sigma^2 I$.

Zauważmy, że jeżeli $X \sim d_n(m, \sigma^2 I)$, to:

$$- E(X_i - m_i) = 0 \text{ dla każdego } i \in \{1, \dots, n\}, \quad (14)$$

$$- D^2 X_i = \sigma^2 \text{ dla każdego } i \in \{1, \dots, n\}, \quad (15)$$

$$- E[(X_i - m_i)(X_j - m_j)] = 0 \text{ dla wszystkich } i, j \in \{1, \dots, n\} \text{ takich, że } i \neq j. \quad (16)$$

Rozważmy wektor losowy \tilde{X} postaci:

$$\tilde{X} = \left(\frac{X_1 - m_1}{\sigma}, \dots, \frac{X_n - m_n}{\sigma} \right).$$

Spełnia on własności (5)–(7).

Istotnie na podstawie własności (14):

$$E(\tilde{X}_i) = E\left(\frac{X_i - m_i}{\sigma}\right) = \frac{1}{\sigma}E(X_i - m_i) \stackrel{(14)}{=} 0 \text{ dla każdego } i \in \{1, \dots, n\}.$$

Dzięki własnościom wariancji zmiennej losowej oraz (15) otrzymujemy:

$$D^2 \tilde{X}_i = D^2 \left(\frac{X_i - m_i}{\sigma} \right) = \frac{1}{\sigma^2} D^2 (X_i - m_i) = \frac{1}{\sigma^2} D^2 X_i \stackrel{(15)}{=} \frac{1}{\sigma^2} \sigma^2 = 1$$

$$\text{dla każdego } i \in \{1, \dots, n\}.$$

W końcu (16) implikuje własność nieskorelowania zmiennych losowych \tilde{X}_i oraz \tilde{X}_j dla wszystkich $i, j \in \{1, \dots, n\}$, takich że $i \neq j$. Istotnie:

$$E(\tilde{X}_i \tilde{X}_j) = E\left[\left(\frac{X_i - m_i}{\sigma} \right) \left(\frac{X_j - m_j}{\sigma} \right) \right] = \frac{1}{\sigma^2} E[(X_i - m_i)(X_j - m_j)] \stackrel{(16)}{=} 0.$$

Na podstawie powyższych ustaleń można stwierdzić, że wektor losowy \tilde{X} ma zerowy wektor wartości oczekiwanych oraz jednostkową macierz kowariancji, czyli $\tilde{X} \sim d_n(0, I)$. Dla wektorów losowych o takich rozkładach nierówność (4) została już wykazana, a więc:

$$\text{Kurt } \bar{X} \geq 1 + (\text{Skew } \bar{X})^2.$$

Zauważmy również, że $\text{Kurt } \bar{X} = \text{Kurt } X$ oraz $(\text{Skew } \bar{X})^2 = (\text{Skew } X)^2$.
Istotnie:

$$\begin{aligned} \text{Kurt } \bar{X} &= \frac{E[(\bar{X} - E\bar{X})^4]}{(D^2 \bar{X})^2} = \frac{E[\bar{X}^4]}{n^2} = \frac{\sum_{i,j=1}^n E[\bar{X}_i^2 \bar{X}_j^2]}{n^2} = \frac{\sum_{i,j=1}^n E\left[\left(\frac{X_i - m_i}{\sigma}\right)^2 \left(\frac{X_j - m_j}{\sigma}\right)^2\right]}{n^2} = \\ &= \frac{\sum_{i,j=1}^n E[(X_i - m_i)^2 (X_j - m_j)^2]}{n^2 \sigma^4} = \frac{E[(X - m)^4]}{(n\sigma^2)^2} = \frac{E[(X - m)^4]}{(D^2 X)^2} = \text{Kurt } X. \end{aligned}$$

Równość dotycząca kwadratów współczynników asymetrii uzasadniają natomiast następujące przekształcenia:

$$\begin{aligned} (\text{Skew } \bar{X})^2 &= \frac{(E[(\bar{X} - E\bar{X})^3])^2}{((D^2 \bar{X})^{\frac{3}{2}})^2} = \frac{(E[\bar{X}^3])^2}{n^3} = \\ &= \frac{\left\langle \left(E\left(\sum_{i=1}^n \bar{X}_i^2 \bar{X}_1\right), \dots, E\left(\sum_{i=1}^n \bar{X}_i^2 \bar{X}_n\right) \right), \left(E\left(\sum_{i=1}^n \bar{X}_i^2 \bar{X}_1\right), \dots, E\left(\sum_{i=1}^n \bar{X}_i^2 \bar{X}_n\right) \right) \right\rangle}{n^3} = \\ &= \frac{\left(E\left(\sum_{i=1}^n \left(\frac{X_i - m_i}{\sigma}\right)^2 \left(\frac{X_1 - m_1}{\sigma}\right)\right), \dots, E\left(\sum_{i=1}^n \left(\frac{X_i - m_i}{\sigma}\right)^2 \left(\frac{X_n - m_n}{\sigma}\right)\right) \right)^2}{n^3} = \\ &= \frac{\frac{1}{\sigma^6} \left(E\left(\sum_{i=1}^n (X_i - m_i)^2 (X_1 - m_1)\right), \dots, E\left(\sum_{i=1}^n (X_i - m_i)^2 (X_n - m_n)\right) \right)^2}{n^3} = \\ &= \frac{(E[X^3])^2}{n^3 \sigma^6} = \frac{(E[X^3])^2}{(n\sigma^2)^3} = \frac{(E[X^3])^2}{(D^2 X)^3} = (\text{Skew } X)^2. \end{aligned}$$

Nierówność (4) została udowodniona.

Literatura

- Bilodeau M., Brenner D. [1999], *Theory of Multivariate Statistics*, Springer-Verlag, New York.
- Budny K. [2009], *Kurtoza wektora losowego*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, nr 76.
- Budny K., Tatar J. [2009], *Kurtosis of a Random Vector – Special Types of Distributions*, „Statistics in Transition – New Series”, vol. 10, nr 3.
- Cramer H. [1958], *Metody matematyczne w statystyce*, PWN, Warszawa.

- Jakubowski J., Sztencel R. [2004], *Wstęp do rachunku prawdopodobieństwa*, Script, Warszawa.
- Kotz S., Johnson N.L., Balakrishnan N. [2000], *Continuous Multivariate Distributions: Model and Applications*, 2nd ed., John Wiley and Sons, New York.
- Mardia K.V. [1970], *Measures of Multivariate Skewness and Kurtosis with Applications*, „Biometrika”, vol. 57, nr 3.
- Tatar J. [1996], *O niektórych miarach rozproszenia rozkładów prawdopodobieństwa*, „Przegląd Statystyczny”, z. 3/4.
- Tatar J. [1999], *Moments of a Random Variable in a Hilbert Space*, „Przegląd Statystyczny”, z. 2.
- Tatar J. [2000], *Asymetria wielowymiarowych rozkładów prawdopodobieństwa*, Materiały z XXXV Konferencji Statystyków, Ekonometryków i Matematyków Akademii Ekonomicznych Polski Południowej zorganizowanej przez Katedrę Statystyki Akademii Ekonomicznej w Krakowie (Osieczany, 23–25 marca 1999 r.), Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.

Some Properties of the Kurtosis of a Random Vector

The paper is a continuation of a discussion of the kurtosis of a random vector. Multivariate kurtosis is defined as the fourth central moment divided by the square of the variance of a random vector. This term is built on the definition of the power of a random vector proposed by J. Tatar. The paper presents selected, essential properties of multivariate kurtosis – among other things the invariance property under a number of affine transformations. Besides that, the relation between kurtosis of the random vector and its skewness is fixed. In view of these properties, the fourth central moment divided by the square of the variance of a random vector may be regarded as a satisfactory measure of multivariate kurtosis.

Keywords: kurtosis of a random vector, power of a vector, multivariate distribution, characteristics of the multivariate distribution.

Marta Targaszewska

Katedra Ekonometrii

Uniwersytet Ekonomiczny we Wrocławiu

Metody pomiaru jakości kształcenia na uczelniach wyższych

Streszczenie

Celem pracy jest przedstawienie wybranych metod pomiaru jakości kształcenia na uczelniach wyższych na podstawie danych marketingowych. Artykuł składa się z dwóch części. W pierwszej w sposób syntetyczny przedstawiono pojęcie jakości i jakości usług edukacyjnych na poziomie wyższym. W drugiej części zaprezentowano metody jej pomiaru z wykorzystaniem modelu SERVQUAL i techniki IPA oraz możliwości aplikacyjnych obu metod na przykładzie Uniwersytetu Ekonomicznego we Wrocławiu.

Słowa kluczowe: jakość usług edukacyjnych, edukacja na poziomie wyższym, SERVQUAL, IPA.

1. Wprowadzenie

Jakość kształcenia na uczelniach wyższych jest jednym z ważniejszym problemów, jakim zajmuje się środowisko akademickie zarówno w Polsce i w Europie, jak i na całym świecie. W Europie ważność problemu podkreślają m.in. deklaracja bolońska, będąca podstawą procesu mającego na celu stworzenie zintegrowanego i konkurencyjnego systemu szkolnictwa wyższego, oraz strategia lizbońska, mająca zapewnić trwałe i zrównoważony wzrost gospodarczy przez inwestycje w kapitał ludzki, a więc w wykształcenie. Jakość jest zatem nie tylko gwarantem podniesienia konkurencyjności jednostki akademickiej, ale ma również wpływ na całe społeczeństwo: przez inwestycje w kapitał ludzki,

przygotowanie absolwentów do sprostania potrzebom rynku pracy, a co za tym idzie – zmniejszenie bezrobocia [*Jakość w badaniach...* 2007, s. 452]. Dlatego też jakość kształcenia na uczelniach wyższych powinna podlegać procesowi zarządzania, który stanowi całościowy systematyzowany dział i procedur dotyczących zapewnienia jakości. Jednym z elementów zarządzania jest ciągła kontrola (pomiar) zjawiska. Celem pracy jest prezentacja dwóch wybranych metod pomiaru jakości usług edukacyjnych – modelu SERVQUAL i techniki IPA, oraz wyników badań uzyskanych z ich wykorzystaniem.

2. Pojęcie jakości usług edukacyjnych

Jakość, zgodnie z definicją zawartą w normach ISO 9000:2000, rozumiana jest jako stopień, w jakim zbiór inherentnych właściwości spełnia wymagania. Postrzegana jest również jako doskonałość, właściwość albo wartość określonego przedmiotu lub zjawiska. Stanowi ona sumę cech wytworu, wyrobu bądź samego działania [Hamrol 2007, s. 19].

Jakość usług kształcenia wyższego jest natomiast definiowana jako poziom, w jakim spełnione są oczekiwania klientów, a więc studentów, ich rodziców, ich przyszłych pracodawców oraz pracowników dydaktycznych [Sztejnberg i Stypułkowska 2005]. Może być ona rozumiana także jako różnica między oczekiwaniami klienta a tym, co otrzymał [Kolman 2009, s. 19]. Jakość usług kształcenia wyższego może być również postrzegana jako stopień spełnienia założonych przez instytucję akademicką celów, które są związane z procesem kształcenia [*Jakość w badaniach...* 2007, s. 452]. Ponadto może być poddana ocenie zewnętrznej, wykorzystującej obiektywne kryteria [Przybylski, Rudnicki i Szwed 2010, s. 22].

Jakość usług edukacyjnych determinowana jest przez następujące czynniki: cechy studenta, nauczanie i uczenie się, wyniki, zasoby oraz kontekst. Każdy z nich posiada składniki szczegółowe. Do składników cech studenta można zaliczyć zdolność, wytrwałość oraz chęć do nauki. Drugą grupę reprezentują stosowane metody nauczania, liczba zajęć, sposób oceny pracy studentów oraz wielkość grup ćwiczeniowych bądź laboratoryjnych. Na wyniki składają się korzyści społeczne, jak również twórcze i emocjonalne umiejętności. Zasoby obejmują środki materialne (materiały dydaktyczne, wydawnictwa dostępne w bibliotece) oraz zasoby ludzkie (kadra naukowo-dydaktyczna, zarządzająca, administracyjna). Ostatni czynnik obejmuje wiele składników, takich jak: warunki ekonomiczne, czynniki społeczne, kulturalne, religijne, wiedzę i umiejętności nauczycieli czy też możliwość współpracy szkoły z organizacjami pozauczelnianymi [Sztejnberg 2008, s. 18–19].

Instytucja akademicka powinna nieustannie polepszać jakość swoich usług. Aby było to możliwe, powinna regularnie dokonywać pomiarów jej poziomu oraz

oceny swojego funkcjonowania na dwóch płaszczyznach [Przybylski, Rudnicki i Szwed 2010, s. 25]: 1) klientów – ocena kompetencji zdobytych w trakcie nauki (przeprowadzana przez studentów i ich rodziców), ocena umiejętności absolwentów danej szkoły wyższej (przeprowadzona przez pracodawców); 2) szkoły wyższej – realizowana przez zewnętrzne instytucje lub samą uczelnię (samodzielnie lub pod kontrolą organizacji zewnętrznej).

3. Modele i techniki pomiaru jakości usług

Pierwszym zaprezentowanym sposobem pomiaru jakości usług edukacyjnych jest model jakości opracowany w 1985 r. przez A. Parasuramana, V.A. Zeithaml i L. Berry'ego [1985, s. 41–50]. Jego ideą jest postrzeganie jakości usługi jako funkcji rozbieżności pomiędzy oczekiwaniami klientów a ich percepcją usługi [Parasuraman, Zeithaml i Berry 1990, s. 23]. Podstawowym założeniem oceny usługi edukacyjnej jest istnienie pięciu luk [Łuczak i Matuszak-Flejszman 2007, s. 343–344]:

- 1) rozbieżności między tym, czego klient oczekuje, a tym, co dostawca usługi – uczelnia, uważa za ważne i satysfakcjonujące dla niego,
- 2) sprzeczności między postrzeganiem przez kierownictwo szkoły (rektor, dziekani) oczekiwań usługobiorców a cechami fizycznymi usługi,
- 3) różnica między wymaganiami jakości usługi a usługą faktycznie dostarczoną,
- 4) sprzeczności między usługą dostarczoną a obiecaną,
- 5) rozbieżności między oczekiwaniami słuchaczy związanymi ze studioowaniem a jego rzeczywistym postrzeganiem; do jej powstania przyczyniają się wszystkie wyżej wymienione luki.

Najczęściej objęciem badań jest ostatnia luka. W pomiarze jej wielkości wykorzystuje się metodę SERVQUAL [Sagan 2003].

Metoda SERVQUAL została opracowana w 1988 r. przez twórców modelu jakości. Uwzględnia ona wykorzystanie dwuczęściowego kwestionariusza badawczego, za pomocą którego odbiorcy przekazują informacje odnoszące się do oczekiwań (część pierwsza) i postrzegania (część druga) usługi aktualnie świadczonej [Szejnberg i Stypułkowska 2005]. Ankieta zawiera pytania (po równej liczbie dla percepcji i oczekiwań) dotyczące obszarów mających najistotniejszy wpływ na jakość usługi (tzw. kryteria jakości). W modelu SERVQUAL podstawowymi kryteriami są: warunki materialne, niezawodność obsługi, odpowiedzialność i chęć współpracy, kompetencje pracowników oraz empatia. W pomiarze jakości usług edukacyjnych na poziomie wyższym warunki materialne stanowią wygląd uczelni i jej pomieszczeń, a także jej zasoby kadrowe i rzeczowe – wyposażenie

sal dydaktycznych. Niezawodność wyraża się przez umiejętność dokładnego, sumiennego i solidnego świadczenia usługi kształcenia – nauczania, jak również dokładne oraz należyte wypełnianie obowiązków przez wykładowców. Odpowiedzialność to szybkość reakcji na oczekiwania i prośby studentów, dostrzeganie przez kadre akademicką problemów, które ich nurtują, oraz gotowość udzielenia im pomocy. Wiedza i umiejętności niezbędne do wykonania usługi, biegłość, umiejętność zdobycia zaufania studentów są wyrazem kompetencji. Ostatnie kryterium – empatia, to uprzejmość, utożsamianie się z potrzebami odbiorców, troska o nich, szacunek i sympatia im okazywane [Bugdol 2008, s. 210].

Odpowiedzi udzielane są za pomocą siedmiostopniowej skali Likerta, zgodnie z którą nota 1 oznacza, że klient zdecydowanie nie zgadza się z przedstawionym stwierdzeniem, a 7 oznacza, że zdecydowanie się z nim zgadza [Choudhury 2011].

Poziom jakości oceniony metodą SERVQUAL można sprowadzić do następującego wzoru [Sagan 2003]:

$$S = \sum_{i=1}^k p_i - e_i, \quad (1)$$

gdzie:

S – subiektywne zadowolenie z usługi,

p_i – ocena postrzegania danego atrybutu usługi i ,

e_i – ocena oczekiwań klienta co do danego atrybutu usługi,

k – liczba pytań.

Subiektywne zadowolenie klienta można interpretować następująco [Sztejnberg, Stypułkowska i Hurek 2005]:

– poziom otrzymanej usługi jest równy poziomowi pożądanemu, gdy $S = 0$,

– postrzegany poziom usługi jest niższy od oczekiwanego, gdy $S < 0$,

– poziom świadczonej usługi jest wyższy od oczekiwanego, gdy $S > 0$.

Wykorzystane narzędzie pomiaru w metodzie SERVQUAL powinno podlegać ocenie rzetelności, a więc badaniu poziomu skorelowania między sobą wymiarów wchodzących w skład kwestionariusza badawczego lub danego kryterium [Sztejnberg, Stypułkowska i Hurek 2005]. Najczęściej stosowaną techniką pomiaru jest współczynnik α -Cronbacha wyrażony wzorem [Brzeziński 2004, s. 475]:

$$\alpha = \frac{k}{k-1} \left[\frac{\text{var}_c - \sum_{i=1}^k \text{var}_i}{\text{var}_c} \right], \quad (2)$$

gdzie:

k – liczba kryteriów w danym kwestionariuszu,

var_c – wariancja całkowita wyników wszystkich kryteriów,

var_i – wariancja i -tego kryterium.

Współczynnik α -Cronbacha przyjmuje wartości z przedziału od 0 do 1; 0 świadczy o braku wyniku prawdziwego, zaś 1 informuje o tym, że wszystkie pozycje są doskonale rzetelne (prawdziwe).

Przedstawiona powyżej skala SERVQUAL umożliwia określenie poziomu zadowolenia klienta. Dzięki niej szkoły wyższe mają możliwość wskazania obszarów wymagających określonych działań naprawczych oraz jasnych standardów jakości. Skala daje również możliwość badania efektywności wprowadzonych zmian przez badanie satysfakcji studentów przed ich wprowadzeniem i po ich wprowadzeniu.

Metodą często stosowaną w pomiarze jakości usługi i satysfakcji klienta jest również metoda *importance-performance analysis* (IPA), którą zaproponowali w 1977 r. J.A. Martilla i J.C. James. Jako narzędzie badawcze wykorzystuje ona kwestionariusz, w którym klienci oceniają według następujących kryteriów: 1) ważności cech usługi uznanych za istotne, 2) postrzegania przez klienta osiągnięć organizacji w zakresie cech kluczowych [Martilla i James 1977, s. 77–79].

Podobnie jak w modelu SERVQUAL odpowiedzi zarówno w przypadku ważności (1 – nieważny, 7 – bardzo ważny), jak i osiągnięć (1 – niezadowolony, 7 – bardzo zadowolony) udzielane są z wykorzystaniem siedmiostopniowej skali Likerta [Sztejnberg 2008].

Otrzymane wyniki – wartości ocen ważności i spostrzeżeń dotyczących osiągnięć w pierwszym kroku – porządkuje się jako: niskie, średnie lub wysokie.

Kolejny krok to ich prezentacja na dwuwymiarowym wykresie, gdzie oś rzędnych odpowiada średniej ważności, zaś oś odciętych średnim osiągnięciom. Punktem przecięcia obu osi jest wartość centralna wyników (najczęściej średnia arytmetyczna lub mediana) ważności i osiągnięć dla wszystkich cech [Kitchroen 2004]. Graficzne przedstawienie wyników kwestionariusza badawczego jest niezbędne do formułowania strategii organizacji akademickiej. Interpretacja wyników zależy od ćwiartki, w której leży wynik dla danej cechy. Ćwiartka pierwsza („kontynuuj dobrą pracę”) obejmuje cechy, które według klientów są ważne przy podejmowaniu decyzji o skorzystaniu z usługi kształcenia wyższego i dla których spostrzeżenia dotyczące osiągnięć są oceniane wysoko. Uczelnia, której cechy są w tej grupie, powinna utrzymywać dotychczasowe wyniki. Cechy znajdujące się w ćwiartce drugiej („skoncentruj się na tym”) są ważne w podejmowaniu decyzji o skorzystaniu z usługi kształcenia na poziomie wyższym, ale ocena osiągnięć instytucji akademickiej nie jest wysoka. Zalecane jest skoncentrowanie się na przymiotach umiejscowionych w tej części matrycy. Cechy znajdujące się w ćwiartce trzeciej („niski priorytet”) reprezentują niski poziom ważności i niski poziom osiągnięć – poniżej średniej. Nie wymagają one jednak bezpośredniej uwagi. Ostatnia, czwarta ćwiartka („możliwa przesada w ulepszaniu”) obejmuje cechy, które są dla klientów mało istotne przy jednoczesnej wysokiej ocenie

dotychczasowych osiągnięć. Zaleceniem dla szkół wyższych mających cechy leżące w tej ćwiartce jest podejmowanie działań w celu zapewnienia ich większej ważności [Sztejnberg 2008, s. 113–114]. Rzetelność pomiaru, podobnie jak w przypadku metody SERVQUAL, mierzy się współczynnikiem α -Cronbacha.

4. Rezultaty przeprowadzonych badań

Aby uzyskać opinię dotyczącą pożądanego i aktualnego poziomu świadczenia usługi edukacyjnej, w pierwszym etapie przeprowadzono ankietę wśród 29 losowo wybranych studentów drugiego roku zarządzania na Wydziale Zarządzania Informatyki i Finansów Uniwersytetu Ekonomicznego we Wrocławiu. Kwestionariusz opracowano zgodnie z procedurą zaproponowaną przez twórców metody SERVQUAL. Obejmował on po 22 stwierdzenia odnoszące się do oczekiwań i percepcji studentów. Pytania dotyczyły każdego z kryterium. W kolejnym etapie zbadano rzetelność zebranych danych zarówno dla całego kwestionariusza, jak i dla każdego kryterium z osobna. Wyniki przedstawiono w tabeli 1.

Tabela 1. Wartości współczynnika rzetelności α -Cronbacha w rozróżnieniu na poszczególne kryteria

Kryterium	Poziom pożądaný	Poziom obserwowany	Poziom według modelu SERVQUAL
Warunki materialne	0,50	0,77	0,68
Niezawodność	0,62	0,80	0,67
Odpowiedzialność	0,69	0,79	0,79
Kompetencje	0,68	0,75	0,77
Empatia	0,89	0,84	0,88
Ogółem	0,91	0,94	0,93

Źródło: opracowanie własne na podstawie zebranych danych.

Na podstawie danych tabeli 1 można wnioskować, że jedynie w przypadku kryterium dotyczącego warunków materialnych dla poziomu pożądanego wartość współczynnika rzetelności jest niewystarczająca (0,50). Dla pozostałych kryteriów wartości współczynnika α -Cronbacha kształtują się powyżej 0,60, co świadczy o jednorodności tych grup stwierdzeń. Ponadto wartości α -Cronbacha dla wszystkich pytań wchodzących w skład kwestionariusza na każdym z poziomów są wysokie i wynoszą ponad 0,90.

Następnie obliczono średnie arytmetyczne dla wyników różnic pomiędzy poziomem obserwowanym a pożądanym. Przyjęto też, że wagi dla każdego kryterium są jednakowe.

Tabela 2. Średnie arytmetyczne dotyczące wyników różnic pomiędzy poziomem obserwowanym a pożądanym dla poszczególnych kryteriów

Kryterium	Średnia
Warunki materialne	-1,95
Niezawodność	-2,31
Odpowiedzialność	-2,40
Kompetencje	-2,30
Empatia	-2,23

Źródło: opracowanie własne na podstawie zebranych danych.

Tabela 3. Wyniki różnic pomiędzy poziomem obserwowanym a pożądanym dla poszczególnych atrybutów

Atrybut	Wynik
Schludny wygląd	-0,66
Przyjazne podejście	-1,69
Dostrzeganie problemów	-2,03
Biegłość w fachu	-2,10
Uprzejmość	-2,17
Rzetelność	-2,17
Atrakcyjne sale dydaktyczne	-2,21
Zaufanie	-2,21
Dokładność i skrupulatność	-2,21
Solidność	-2,28
Sumienność	-2,31
Otwartość	-2,31
Życzliwość	-2,34
Skuteczność	-2,38
Wyposażenie sal	-2,41
Okazywanie uwagi	-2,41
Pomoc	-2,45
Obowiązkowość	-2,48
Pomoce naukowe w salach dydaktycznych	-2,52
Zrozumienie potrzeb	-2,52
Szybkość reakcji	-2,79

Źródło: opracowanie własne na podstawie zebranych danych.

Analizując dane tabeli 2, można stwierdzić, że najwyżej oceniono warunki materialne, ale ich pomiar był nierzetelny. W przypadku pozostałych kryteriów wyniki nieznacznie się od siebie różniły i mieściły się w przedziale od $-2,40$ do $-2,23$. Ogólny wynik dla wszystkich pytań ukształtował się na poziomie $-2,24$. Rezultaty dla poszczególnych atrybutów przedstawia tabela 3, w której wyniki uporządkowano według malejącego poziomu satysfakcji.

Na podstawie danych tabeli 3 można stwierdzić, że żaden z atrybutów nie spełnił oczekiwań studentów biorących udział w badaniach. Najniższej oceniona została szybkość reakcji prowadzących zajęcia, najwyżej zaś ich schludny wygląd oraz przyjazne podejście. Wysoko ocenione zostały również uprzejmość oraz dostrzeganie problemów przez nauczycieli akademickich, a także biegłość w fachu i rzetelność.

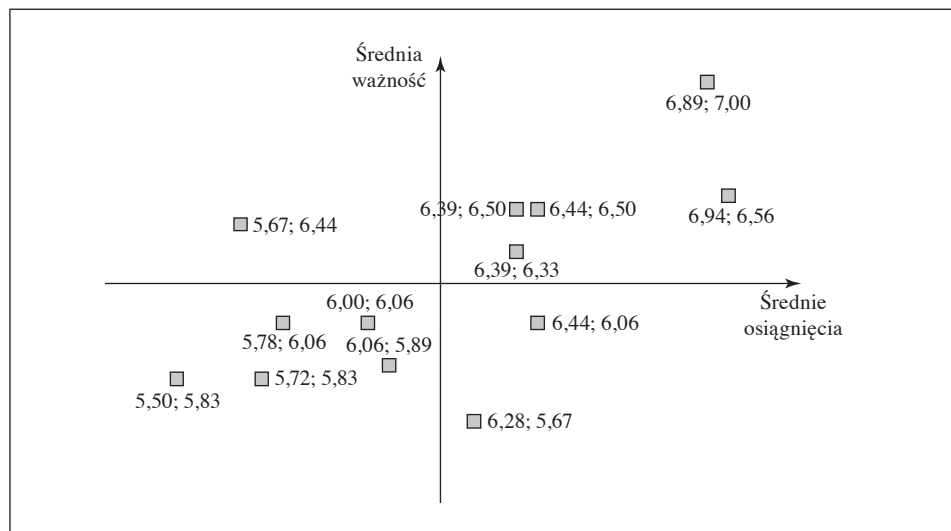
W celu rozpoznania słabych i mocnych stron organizowanych zajęć na Uniwersytecie Ekonomicznym we Wrocławiu przepytano 17 losowo wybranych studentów drugiego roku Wydziału Zarządzania, Informatyki i Finansów oraz uczestniczących w zajęciach prowadzonych przez tego samego nauczyciela akademickiego. Studenci udzielali odpowiedzi na 26 pytań dotyczących ważności określonych atrybutów i ich osiągnięć. Podobnie jak w modelu SERVQUAL zebrane dane poddano analizie rzetelności. Obliczono współczynnik α -Cronbacha, który dla pytań dotyczących ważności i osiągnięć wyniósł odpowiednio 0,65 i 0,81. Następnie oszacowano średnie arytmetyczne dotyczące wszystkich aspektów, do których odnosiły się pytania zawarte w kwestionariuszu. Atrybuty oraz wyniki badań w podziale na ważność i osiągnięcia przedstawia tabela 4.

Tabela 4. Wyniki analizy IPA dla poszczególnych atrybutów

Atrybut	Osiągnięcia	Ważność
Zrozumiały cel przedmiotu	6,00	6,06
Zrozumiała treść kształcenia	6,28	5,67
Jasno postawione wymagania	6,39	6,50
Spójność zajęć	5,78	6,06
Zrozumiały materiał	5,67	6,44
Uczestniczenie w zajęciach z entuzjazmem	5,50	5,83
Ciekawe treści	5,72	5,83
Materiał prezentowany w jasny sposób	6,44	6,50
Logiczne uporządkowanie poruszanych zagadnień	6,44	6,06
Odpowiednia ilość materiału	6,39	6,33
Motywowanie do pracy	6,06	5,89
Wiedza prowadzącego	6,89	7,00
Pomoc ze strony prowadzącego	6,94	6,56

Źródło: opracowanie własne na podstawie zebranych danych oraz [Lewis 2004].

Otrzymane wyniki zostały również rzutowane na układ współrzędnych (rys. 1), gdzie punkt przecięcia osi wyznaczono na podstawie średnich arytmetycznych dla osiągnięć i ważności. Wartości średnich wyniosły odpowiednio 6,19 i 6,21.



Rys. 1. Siatka IPA

Źródło: opracowanie własne na podstawie zebranych danych.

Jak wynika z analizy rys. 1, w pierwszej ćwiartce („kontynuuj dobrą pracę”) znajdują się takie atrybuty, jak: wiedza prowadzącego, jasny sposób prezentowania materiału, pomoc ze strony prowadzącego, jasno postawione wymagania, a także odpowiednia ilość materiału. Są one ważne dla studentów oraz jednocześnie wysoko przez nich oceniane. Równie wysoko oceniane są logiczne uporządkowanie poruszanych zagadnień i zrozumiała treść kształcenia, ale ich ważność jest oceniana nisko i nie wymagają one bezpośredniej uwagi. Obserwacji nie wymagają również takie mało ważne atrybuty, jak: rozumiały cel przedmiotu, spójność zajęć, uczestniczenie w zajęciach z entuzjazmem, ciekawe treści przedmiotu oraz motywowanie do pracy. Najwięcej uwagi i pracy wymaga cecha znajdująca się w ćwiartce drugiej – zrozumienie materiału, i to na niej powinien się koncentrować prowadzący zajęcia.

5. Zakończenie

Zastosowane metody pozwoliły ocenić satysfakcję wybranych studentów Uniwersytetu Ekonomicznego we Wrocławiu oraz wskazać mocne i słabe strony

badanej uczelni. Pierwsza zaprezentowana metoda pozwoliła na wykazanie wielkości luki pomiędzy obserwowanym a pożądanym poziomem cech mających wpływ na jakość świadczonych usług. Z analizy wynika, że na poziomie zarówno kryteriów, jak i pojedynczych pytań oczekiwania przewyższają aktualną ocenę określonych atrybutów. Druga metoda pozwoliła na ujawnienie, nad którymi z aspektów decydujących o jakości usługi należy pracować, a które są wystarczająco dobrze postrzegane przez studentów. W obu badaniach wysoko oceniono wiedzę prowadzących oraz ich rzetelność. Najślabiej oceniono szybkość reakcji nauczycieli akademickich na prośby studentów, która może mieć istotny wpływ na trudności ze zrozumieniem materiału, co wykazano za pomocą techniki IPA. Są to więc atrybuty, nad którymi trzeba pracować, aby zwiększać satysfakcję klientów badanej uczelni.

Pomiar jakości kształcenia w zakresie szkolnictwa wyższego na podstawie danych marketingowych jest bardzo złożonym procesem. Przedstawione badania są przykładem pomiaru satysfakcji jednej z grup klientów uczelni – studentów, ale powinno się również analizować satysfakcję innych grup klientów szkół wyższych (studentów i ich rodziców, przyszłych pracodawców oraz nauczycieli akademickich). Analizą powinni zostać objęci także absolwenci danej uczelni, którzy mogą ocenić efekty kształcenia i przydatność zdobytej wiedzy i umiejętności.

Literatura

- Brzeziński J. [2004], *Metodologia badań psychologicznych*, wyd. 5, Wydawnictwo Naukowe PWN, Warszawa.
- Bugdol M. [2008], *Zarządzanie jakością w urzędach administracji publicznej. Teoria i praktyka*, wyd. 1, Difin, Warszawa.
- Choudhury K. [2011], *A Model of Service Quality in the Education Service*, www.iimcal.ac.in/research/Download/education-Koushiki.pdf (dostęp: 15.01.2011).
- Hamrol A. [2007], *Zarządzanie jakością z przykładami*, wyd. 2 zm., Wydawnictwo Naukowe PWN, Warszawa.
- Jakość w badaniach i dydaktyce szkół wyższych*, red. S. Doroszewicz i A. Kobylińska [2007], Oficyna Wydawnicza SGH, Warszawa.
- Kitchroen K. [2004], *Literature Review: Service Quality in Educational Institutions*, „ABAC Journal”, vol. 24, nr 2, www.journal.au.edu/abac_journal/2004/may04/abacvol24no2_artical02.pdf (dostęp: 7.01.2011).
- Kolman R. [2009], *Kwalitologia – wiedza o różnych dziedzinach jakości*, wyd. 1, Wydawnictwo Placet, Warszawa.
- Lewis R. [2004], *Importance-Performance Analysis*, „Australasian Journal of Engineering”, vol. 2, www.aae.com.au/journal/2004/lewis04.pdf (dostęp: 12.02.2011).
- Łuczak J., Matuszak-Flejszman A. [2007], *Metody i techniki zarządzania jakością*, wyd. 1, Wydawnictwo Quality Progress, Poznań.
- Martilla J.A., James J.C. [1977], *Importance-Performance Analysis*, „Journal of Marketing”, vol. 41.

- Parasuraman A., Zeithaml V.A., Berry L.L. [1985], *A Conceptual Model of Service Quality and Its Implications for Future Research*, „Journal of Marketing”, vol. 49.
- Parasuraman A., Zeithaml V. A., Berry L.L. [1990], *Delivering Quality Service. Balancing Customer Perceptions and Expectations*, 1st ed., The Free Press, New York.
- Przybylski W., Rudnicki S., Szwed A. [2010], *Ewaluacja jakości dydaktyki w szkolnictwie wyższym. Metody, narzędzia, dobre praktyki*, wyd. 1, Wydawnictwo Wyższej Szkoły Europejskiej, Kraków.
- Sagan A. [2003], *Skale jako podstawowy instrument pomiaru w badaniach satysfakcji i lojalności*, www.statsoft.pl/czytelnia/marketing/skale.pdf (dostęp: 2.12.2010).
- Sztejnberg A. [2008], *Doskonalenie usług edukacyjnych. Podstawy pomiaru jakości kształcenia*, wyd. 1, Wydawnictwo Uniwersytetu Opolskiego, Opole.
- Sztejnberg A., Stypułkowska M. [2005], *Ocena względnej ważności wymiarów SERVQUAL jako metoda wspomaganie jakości kształcenia w szkole wyższej [w:] Media i metody wspomagające jakość kształcenia*, red. K. Jankowski, B. Sitarska i C. Tkaczuk, Wydawnictwo Akademii Podlaskiej, Siedlce, www.inp.uni.opole.pl/~sztejnberg/komspol/ocena.pdf (dostęp: 27.01.2011).
- Sztejnberg A., Stypułkowska M., Hurek J. [2005], *Zastosowanie modelu SERVQUAL w mierzeniu jakości usług edukacyjnych [w:] Aktualni otázky výuky chemie / Actual Questions of Chemistry Education*, red. M. Bilek, Hradec Králové, www.inp.uni.opole.pl/~sztejnberg/komspol/servqual.pdf (dostęp: 27.01.2011).

Methods for Measuring Quality in Higher Education

The aim of this article is to present marketing methods of measuring the quality of higher education. The article consists of two parts. The first concerns theoretical issues about quality and quality service of higher education. The second presents two methods of measuring quality service in higher education (SERVQUAL, IPA).

Keywords: quality in higher education, higher education, SERVQUAL, IPA.

Jacek Stelmach

Katedra Statystyki

Uniwersytet Ekonomiczny w Katowicach

O porównaniu dwóch populacji

Streszczenie

Porównywanie dwóch populacji jest interesującym zagadnieniem statystycznym. Dotyczy znajdowania istotnych statystycznie różnic na podstawie pozyskanych prób.

Najczęściej sprawdzane są hipotezy o równości pewnego charakterystycznego parametru: wartości średniej, wariancji lub frakcji. Najskuteczniejsze parametryczne testy wymagają spełnienia założenia o zgodności rozkładów badanych populacji z rozkładem normalnym. Istnieją jednak przypadki, w których kluczowe znaczenie może mieć porównanie kształtu populacji wielowymiarowych. Dodatkowo rozkłady badanych populacji są nieznane lub też nie mogą być uznane za rozkłady normalne wielowymiarowe.

Niniejszy artykuł przedstawia wyniki badań dotyczących weryfikacji hipotezy statystycznej o braku różnic pomiędzy populacjami wykorzystującej badanie różnic pomiędzy wektorami własnymi. Statystyki testowe zawierające różnice pomiędzy wektorami własnymi badanych populacji pozwalają na badanie różnic w kształcie populacji niezależnie od ich wartości średnich lub wariancji. Mogą więc być wykorzystane do testowania zmienności zjawisk w czasie nawet w obliczu trendu. Zaproponowano weryfikację hipotez statystycznych za pomocą testów permutacyjnych, co zwalnia z konieczności badania zgodności z rozkładem normalnym oraz pozwala na stosowanie różnych statystyk testowych. W podsumowaniu dokonano oceny własności proponowanych testów z wykorzystaniem metody Monte Carlo.

Słowa kluczowe: testy statystyczne, testy permutacyjne, wektory własne, metoda Monte Carlo.

1. Wprowadzenie

Porównanie dwóch populacji stanowi interesujące zagadnienie statystyczne. Ma ono także duże zastosowanie praktyczne, ponieważ umożliwia porównywanie

dwóch produktów, usług, procesów przemysłowych czy zjawisk gospodarczych. Umożliwia także weryfikację zmienności tych zjawisk w czasie lub ich wrażliwości na określony czynnik (bądź grupę czynników). Z reguły badane urządzenie czy zjawisko opisuje zestaw parametrów tworzący zmienną wielowymiarową. Prowadzi to do analizy wielowymiarowej wyników obserwacji pochodzących z dwóch populacji. Testowanie zgodności może polegać na porównaniu [Domański i Pruska 2000, s. 167]:

- 1) gęstości,
- 2) dystrybuant,
- 3) momentów,
- 4) statystyk pozycyjnych.

Najczęściej stosuje się porównywanie dwóch populacji ze względu na pewien parametr, np. wartość średnią czy wariancję. Najbardziej popularne metody (test T^2 Hotellinga, analiza MANOVA) służące do weryfikacji hipotez o równości parametrów wymagają jednak spełnienia założeń o:

- rozkładach obserwacji zbliżonych do wielowymiarowego rozkładu normalnego,
- macierzach kowariancji tych obserwacji jednakowych i niezmiennych,
- niezależności obserwacji.

Dość często jednak mamy do czynienia ze zjawiskami, które nie spełniają powyższych założeń, bądź też nie potrafimy tego z zadowalającym stopniem pewności potwierdzić. P.K. Ito [1980, s. 220] twierdzi, że metoda MANOVA jest odporna na odstępstwa od założenia o normalności i stałości macierzy kowariancji, ale dodaje, że nie dotyczy to przypadków, w których licznosci prób są małe i różne od siebie. Przypadki takie są w rzeczywistości dość częste. Wtedy wnioskowanie statystyczne może opierać się na testach nieparametrycznych. Często jednak tablice wartości krytycznych tych testów, które zostały opublikowane, dotyczą tylko niewielkiej liczby wymiarów [Domański i Pruska 2000, s. 184]. Ograniczenie, jakim jest brak tablic wartości krytycznych statystyk testowych, można ominąć, stosując testy permutacyjne. Popularną statystyką testową w weryfikacji hipotezy o równości wartości średnich prób wielowymiarowych jest np. statystyka, którą podaje P.I. Good [2005, s. 174]:

$$\mathbf{T}^2 = (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2)^T \mathbf{C}^{-1} (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2), \quad (1)$$

gdzie:

$\mathbf{X}_1 = (x_{11}, x_{12}, \dots, x_{1m})$, $\mathbf{X}_2 = (x_{21}, x_{22}, \dots, x_{2n})$ – macierze obserwacji wielowymiarowych,

N – liczba permutacji,

k – ilość wymiarów,

\mathbf{C} – macierz elementów:

$$c_{ij} = \sum_{l=1}^k \sum_{m=1}^{n_k} (X_{lmi} - U_i)(X_{lmj} - U_j), \quad (2)$$

$$U_j = N^{-1} \sum_{l=1}^k \sum_{m=1}^{n_k} X_{lij}. \quad (3)$$

Metoda omawiana w artykule dotyczy badania miar kształtu, wykorzystuje właściwości wektorów własnych w analizie głównych składowych (PCA), w której weryfikację hipotezy statystycznej przeprowadza się za pomocą testów permutacyjnych. Pozwala to nie tylko na porównanie dwóch populacji dotyczące różnic pomiędzy wartościami średnimi lub rozproszeniem, ale także na badanie różnic w kształcie populacji niezależnie od ich wartości średnich lub wariancji. Sama istota testów permutacyjnych zwalnia z konieczności weryfikacji założeń o postaci badanych rozkładów.

2. Prezentacja hipotezy badawczej

Opis problemu

Powodem podjęcia opisanych w niniejszym opracowaniu badań była konieczność porównania dwóch serii produkcyjnych czujników przemysłowych. Ze względu na szeroki dopuszczalny przedział regulacji parametrów tych czujników dla różnych serii produkcyjnych nie można było się oprzeć na typowych testach różnicy wartości średnich lub wariancji. Ważne było natomiast określenie, czy skupiska wielowymiarowe obrazujące zestawy parametrów elektrycznych tych czujników są do siebie podobne. Zrodził się pomysł wykorzystania jako pewnego wskaźnika kształtu tych skupisk zestawów wektorów własnych. Zauważono, że jeśli wektory własne obliczone dla pozyskanych dwóch prób byłyby do siebie zbliżone, brak byłoby podstaw do odrzucenia hipotezy o zgodności testowanych serii produkcyjnych. Nieznajomość postaci rozkładów (asymetryczny, ucięty), a w szczególności wartości krytycznych stała się przyczyną wykorzystania w eksperymencie testów permutacyjnych.

Postawienie hipotezy

Przyjmijmy, że zmienne losowe x_1, \dots, x_k będą próbą wylosowaną z badanej populacji o rozkładzie określonym dystrybuantą $F_1(x)$, zaś zmienne losowe y_1, \dots, y_m będą próbą wylosowaną z badanej populacji o rozkładzie określonym dystrybuantą $F_2(y)$. Wtedy hipoteza badawcza przyjmie postać:

$$H_0: F_1(x) = F_2(y). \quad (4)$$

3. Założenia teoretyczne – analiza głównych składowych (PCA)

Analiza głównych składowych służy dekompozycji zmienności próby wielowymiarowej reprezentowanej wektorem \mathbf{X} na zbiór składowych, gdzie:

- każda składowa jest ortogonalna względem pozostałych,
- pierwsza składowa wyjaśnia największą część wariancji zmiennych X_i ,
- kolejne składowe – największą część pozostałej wariancji itd.

Główne składowe Y_i stanowią kombinację liniową pierwotnych zmiennych X_i :

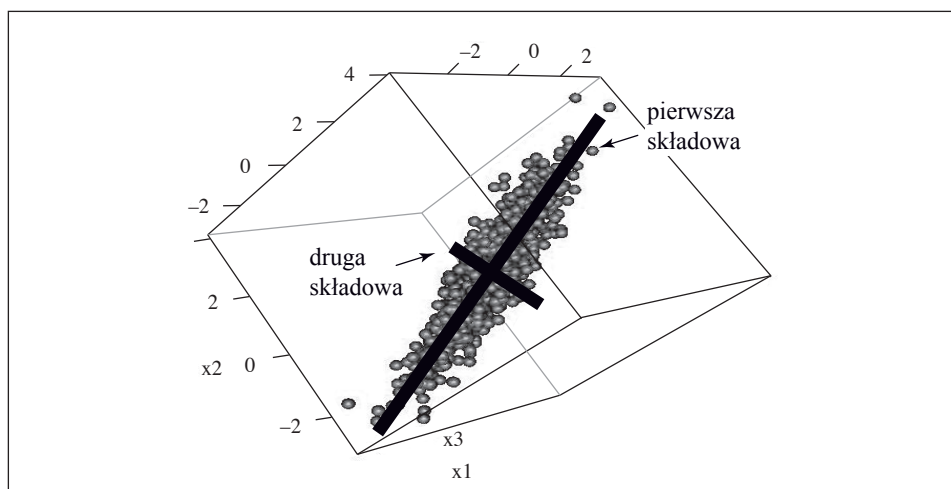
$$Y_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ik}X_k, \quad (5)$$

Geometrycznie główne składowe można interpretować jako kierunki w przestrzeni wielowymiarowej, w której próba jest najbardziej rozproszona (rys. 1). Kierunki te określają kolejne wektory własne macierzy kowariancji uszeregowane według malejących wartości własnych. Analiza w interesującym nas zakresie sprowadza się do znalezienia macierzy wektorów własnych \mathbf{V} oraz wektora wartości własnych $\boldsymbol{\lambda}$, dla macierzy kowariancji $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = \frac{1}{N-1} \mathbf{X}\mathbf{X}^T. \quad (6)$$

Matematyczny opis modelu opiera się na równaniu charakterystycznym:

$$\boldsymbol{\Sigma}\mathbf{V} = \boldsymbol{\lambda}\mathbf{V}. \quad (7)$$



Rys. 1. Interpretacja geometryczna głównych składowych

Źródło: opracowanie własne.

Wektor wartości własnych określa ilość wariacji wyjaśnianych przez poszczególne składowe główne. Jednym z częściej stosowanych kryteriów określenia liczby wykorzystywanych składowych głównych jest – stosowane dalej w rozważaniach – kryterium wartości własnych, w którym pozostawia się te składowe główne, dla których wartości własne są większe od 1.

4. Testy permutacyjne

Testy permutacyjne należą do metod statystycznych, których praktyczne wykorzystanie w ostatnich latach ogromnie wzrasta. Jest to wynik wzrostu mocy obliczeniowej stosowanych komputerów.

W testach permutacyjnych nie jest konieczne spełnienie założeń o normalności rozkładów badanych zjawisk, a weryfikacja hipotez jest możliwa nawet dla prób o niewielkich licznosciach. Istotą testów permutacyjnych jest utworzenie własnej, najlepiej dobranej do konkretnego przypadku statystyki testowej, a następnie symulacyjne przybliżenie jej rozkładu – jako rozkładu statystyk obliczonych dla wszystkich możliwych permutacji. Najczęściej ze względu na ogromną liczbę permutacji stosuje się losowania bezzwrotne wystarczającej ilości kombinacji. Dokładność tej metody jest dla większości przypadków wystarczająca już dla 1000 permutacji [Hesterberg i in. 2003, s. 45]. Na podstawie otrzymanych wyników wyznacza się empiryczny rozkład wybranej statystyki testowej. Stwierdzamy brak podstaw do odrzucenia hipotezy H_0 , jeśli spełniony jest warunek:

$$T_{N,\alpha/2} \leq T_0 \leq T_{N,1-\alpha/2}, \quad (8)$$

gdzie: $T_{N,\alpha/2}, T_{N,1-\alpha/2}$ są kwantylami empirycznego rozkładu statystyki T_N rzędu $\alpha/2$ oraz $1 - \alpha/2$. Jeżeli warunek (8) nie jest spełniony, to odrzucamy hipotezę H_0 .

5. Opis eksperymentu

Ze względu na brak doświadczenia oraz znanych opisów literaturowych tego typu podejścia do porównania dwóch populacji w eksperymencie przeprowadzono badania dla całego szeregu prób 5-wymiarowych otrzymanych drogą symulacji komputerowych. Celem badań było poznanie własności i ograniczeń tej metody statystycznej. W każdej z symulacji pobierano metodą Monte Carlo (100 losowań) dwie podpróby z prób o kontrolowanych typach rozkładów. Licznosci podprób wynosiły odpowiednio:

- 10 i 10 obserwacji,
- 30 i 30 obserwacji,

- 50 i 50 obserwacji,
- 200 i 200 obserwacji,
- 10 i 50 obserwacji,
- 10 i 200 obserwacji.

W analizach rozważano próby o symulowanych rozkładach, tworząc następujące przypadki testowe.

Przypadek testowy 1

Pierwsza próba – rozkład normalny wielowymiarowy o wektorze wartości średnich $[0, 0, 0, 0, 0]$ oraz jednostkowej macierzy kowariancji.

Druga próba – rozkład normalny wielowymiarowy o wektorze wartości średnich przesuwany od $[0, 0, 0, 0, 0]$ do $[x, x, x, x, x]$ (gdzie $x = 0,0; 0,1; \dots; 1,0$) oraz jednostkowej macierzy kowariancji.

Przypadek testowy 2

Pierwsza próba – rozkład normalny wielowymiarowy o wektorze wartości średnich $[0, 0, 0, 0, 0]$ oraz jednostkowej macierzy kowariancji.

Druga próba - rozkład normalny wielowymiarowy o wektorze wartości średnich $[0, 0, 0, 0, 0]$ oraz macierzy kowariancji postaci jak w tabeli 1 (gdzie $x = 0,0; 0,1; \dots; 0,9$).

Tabela 1. Macierz kowariancji dla przypadku 2

		C_1			
C_2	1	x	0	0	0
	x	1	0	0	0
	0	0	1	0	0
	0	0	0	1	0
	0	0	0	0	1

Źródło: opracowanie własne.

Przypadek testowy 3

Pierwsza próba – rozkład normalny wielowymiarowy o wektorze wartości średnich $[0, 0, 0, 0, 0]$ oraz jednostkowej macierzy kowariancji.

Druga próba – rozkład normalny wielowymiarowy o wektorze wartości średnich $[0, 0, 0, 0, 0]$ oraz macierzy kowariancji postaci jak w tabeli 2 (gdzie $x = 0,0; 0,1; \dots; 0,9$):

Tabela 2. Macierz kowariancji dla przypadku 3

	C_1				
C_2	1	x	x	0	0
	x	1	x	0	0
	x	x	1	0	0
	0	0	0	1	0
	0	0	0	0	1

Źródło: opracowanie własne.

Przypadek testowy 4

Pierwsza próba – rozkład normalny wielowymiarowy o wektorze wartości średnich $[0, 0, 0, 0, 0]$ oraz jednostkowej macierzy kowariancji.

Druga próba – rozkład wielowymiarowy odkształcony od rozkładu normalnego (takie same odkształcenie dla każdego wymiaru), standaryzowany, o miarach asymetrii i spłaszczenia jak w tabeli 3.

Tabela 3. Miary skośności i spłaszczenia dla przypadku 4

Wariant	Skośność	Kurtoza
1	0,5	0,25
2	0,01	0,8
3	0,15	0,4

Źródło: opracowanie własne.

Przypadek testowy 5

Pierwsza próba i druga próba – rozkłady normalne wielowymiarowe o wektorze wartości średnich $[0, 0, 0, 0, 0]$ oraz macierzach kowariancji, odpowiednio (gdzie $x = 0,0; 0,1; \dots; 0,9$). Macierz kowariancji dla przypadku 5 (pierwsza i druga próba) przedstawiono w tabelach 4 i 5.

Tabela 4. Macierz kowariancji dla przypadku 5 – pierwsza próba

	C_1				
C_2	1	x	x	0	0
	x	1	x	0	0
	x	x	1	0	0
	0	0	0	1	0
	0	0	0	0	1

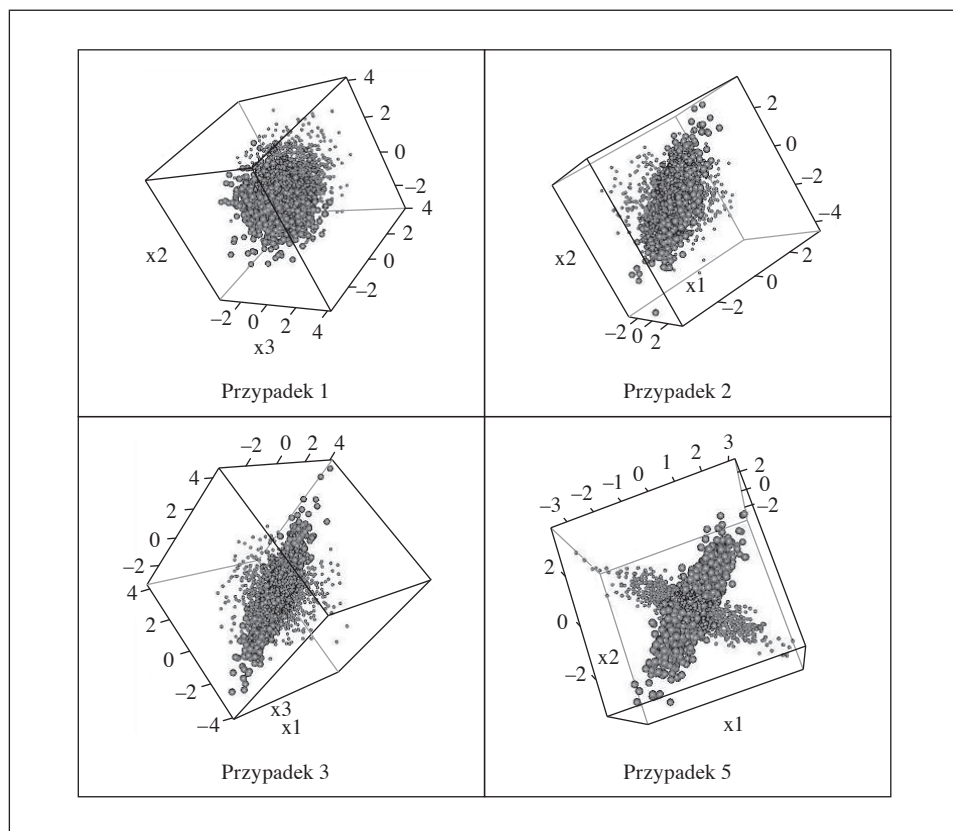
Źródło: opracowanie własne.

Tabela 5. Macierz kowariancji dla przypadku 5 – druga próba

		C_1			
C_2	1	$-x$	$-x$	0	0
	$-x$	1	$-x$	0	0
	$-x$	$-x$	1	0	0
	0	0	0	1	0
	0	0	0	0	1

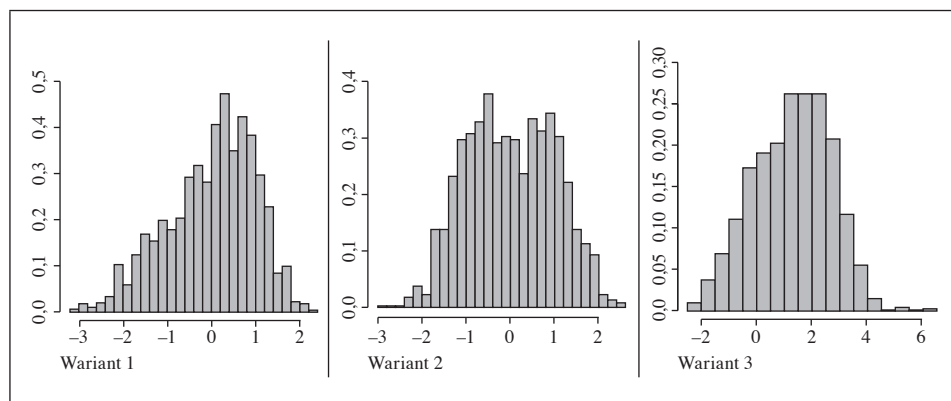
Źródło: opracowanie własne.

Na rys. 2 przedstawiono kształty elipsoid przypadków 1–3 oraz 5 w widoku 3D, a na rys. 3 histogramy jednej ze zmiennych przypadku 4 dla wszystkich określonych wariantów.



Rys. 2. Obraz 3D elipsoid reprezentujących przypadki testowe 1–3 i 5

Źródło: opracowanie własne.



Rys. 3. Przykładowe histogramy jednego z wymiarów dla wariantów przypadku 4
Źródło: opracowanie własne.

Istotne znaczenie w stosowaniu testów permutacyjnych ma dobór statystyki testowej, dość często dobiera się statystyki testowe zbliżone bądź identyczne ze statystykami testów parametrycznych. W tym przypadku ze względu na brak takich testów parametrycznych badania przeprowadzono dla dwóch statystyk (ST_1 i ST_2):

Statystyka ST_1

Stanowi ją długość (w metryce euklidesowej) wektora będącego różnicą pierwszych wektorów własnych:

$$ST_1 = \|\Delta \mathbf{V}_1\| = \mathbf{V}_{11} - \mathbf{V}_{21} = \sqrt{\sum_{i=1}^k (v_{11i} - v_{21i})^2}, \quad (9)$$

gdzie:

$\mathbf{V}_{11}, \mathbf{V}_{21}$ – pierwsze wektory własne pierwszej i drugiej próby,
 k – liczba wymiarów.

Statystyka ST_2

Stanowi ją suma ważona długości (w metryce euklidesowej) wektorów będących różnicami kolejnych wektorów własnych – zgodnie z kryterium wartości własnych (a więc aż do tych wektorów własnych, dla których odpowiednio obie wartości własne są większe od 1). Wagi stosowanymi w próbie są średnie arytmetyczne proporcji wariacji wyjaśnionych przez odpowiednią główną składową. Taka statystyka umożliwia lepsze odwzorowanie kształtu elipsoid wielowymiarowych, szczególnie kiedy więcej wartości własnych ma wartości większe od 1:

$$ST_2 = \sum_{j=1}^m \|\mathbf{V}_{1j} - \mathbf{V}_{2j}\| \frac{\lambda_{1j} + \lambda_{2j}}{2k}, \quad \text{jeśli } \lambda_{1j}, \lambda_{2j} > 1, \quad (10)$$

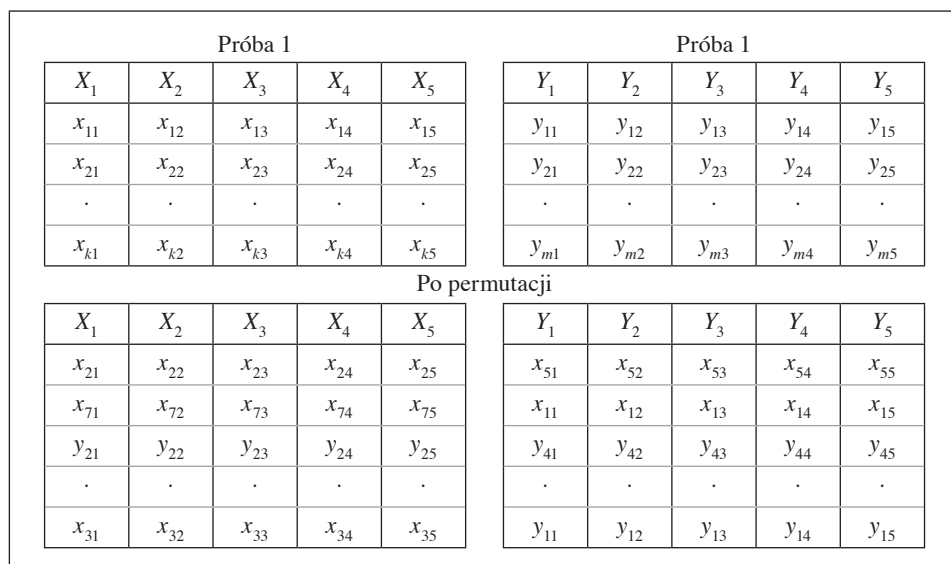
gdzie:

$\lambda_{1j}, \lambda_{2j}$ – j -te wartości własne pierwszej i drugiej podpróby,
 m – liczba głównych składowych spełniających kryterium wartości własnych.

Test permutacyjny przeprowadzono $N = 1000$ razy zgodnie z przedstawioną poniżej sekwencją.

1. Po wybraniu odpowiedniej statystyki obliczyć jej wartość dla badanej próby – T^* .

Dokonać permutacji danych, która niszczy zależności pomiędzy zmiennymi oraz tworzy – N razy – dwie grupy zmiennych, zgodnie z rys. 4.



Rys. 4. Ilustracja tworzenia kolejnych permutacji prób

Źródło: opracowanie własne.

2. Obliczyć wartość wybranej statystyki testowej dla wylosowanych grup – T_i ,
 $i = 1, 2, \dots, N$.

3. Oszacować wartość p -value jako ASL (obszar krytyczny prawostronny):

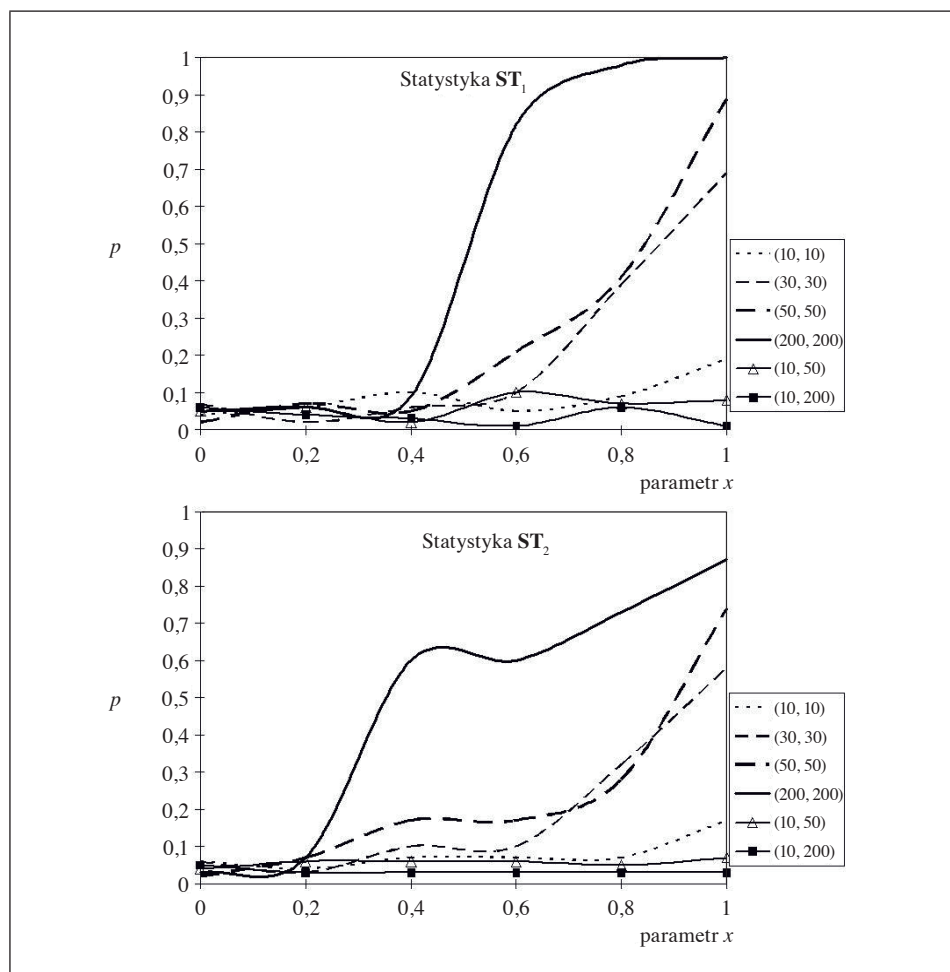
$$ASL = \frac{\text{card}(T^* > T_i)}{M}. \quad (11)$$

Jeśli oszacowana wielkość ASL jest większa od przyjętego poziomu istotności α , nie ma podstaw do odrzucenia testowanej hipotezy.

6. Wyniki obliczeń

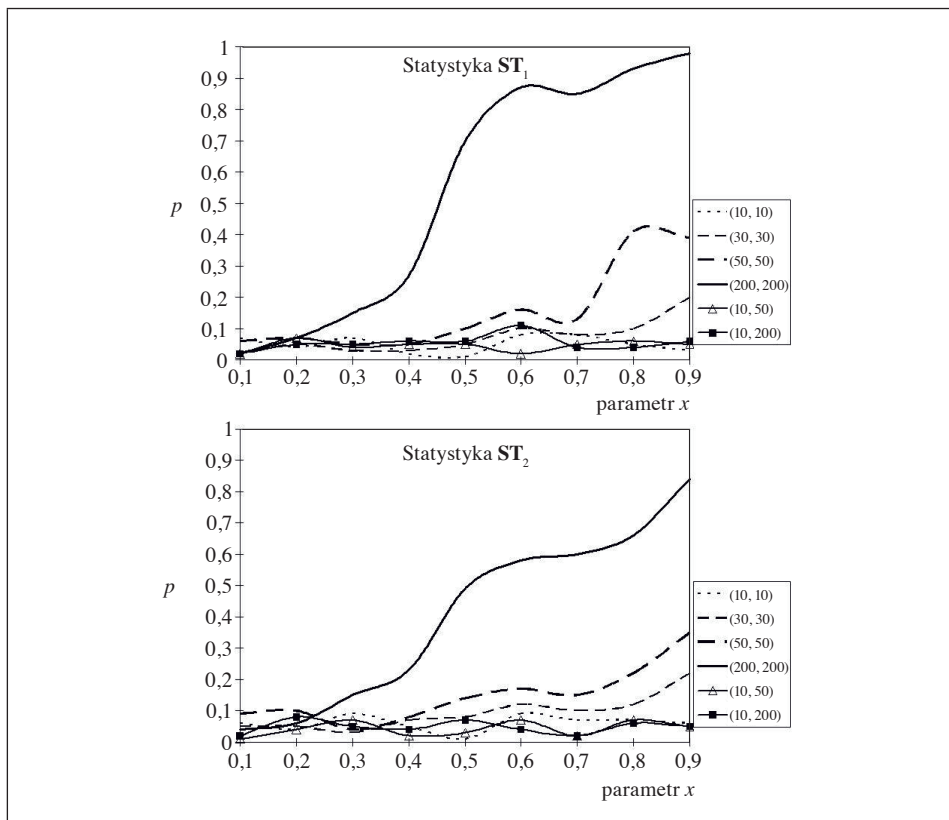
Dla każdego z przypadków na podstawie 100 symulacji Monte Carlo o licznosciach podprób określonych w punkcie 5 obliczono prawdopodobieństwo odrzucenia hipotezy H_0 dla poziomów istotności $\alpha = 0,05$ oraz $\alpha = 0,10$.

Wyniki przedstawiono na rys. 5–8, na których wykreślono krzywe prawdopodobieństwa odrzucenia H_0 w zależności od parametru x oraz licznosci podprób dla przypadków testowych 1–3 i 5 oraz w tabeli 6 (przypadek testowy 4).



Rys. 5. Przypadek testowy 1

Źródło: obliczenia własne.



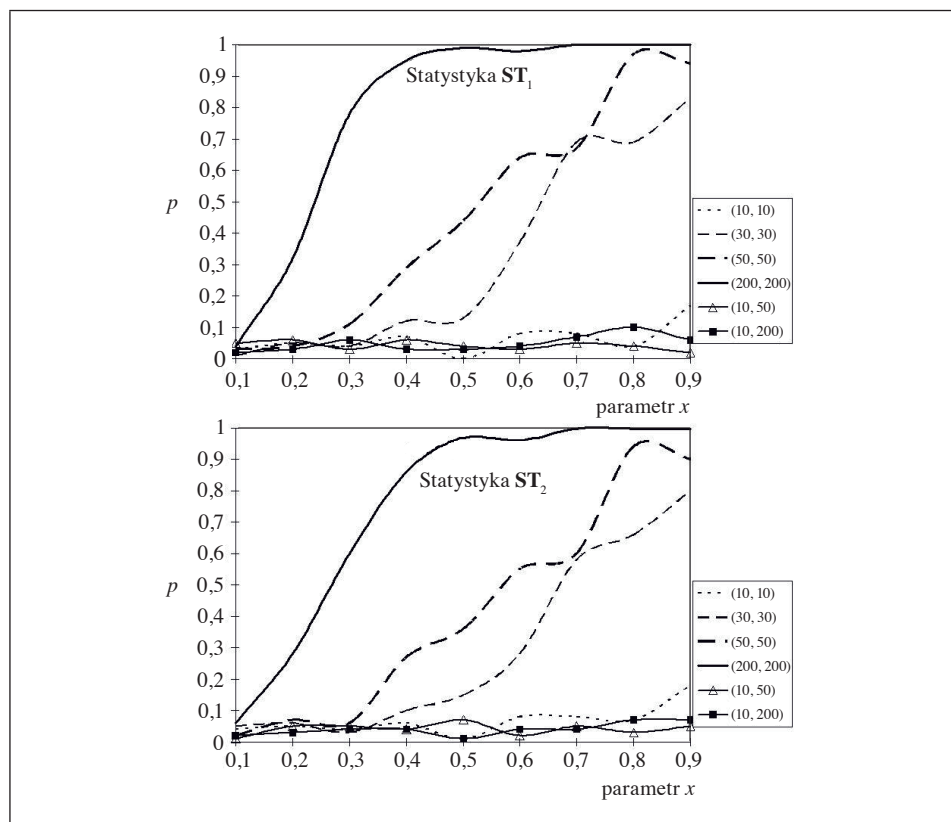
Rys. 6. Przypadek testowy 2

Źródło: obliczenia własne.

Tabela 6. Prawdopodobieństwa odrzucenia H_0 dla wszystkich wariantów przypadku testowego 4 w zależności od liczności próbek

Wariant testowy	Liczności próbek (n_1, n_2)					
	(10, 10)	(30, 30)	(50, 50)	(200, 200)	(10, 50)	(10, 200)
Statystyka ST_1						
1	0,31	0,96	1	1	0,09	0,08
2	0,46	1	1	1	0,1	0,08
3	0,16	0,78	0,94	1	0,09	0,05
Statystyka ST_2						
1	0,31	0,96	1	1	0,09	0,08
2	0,46	1	1	1	0,1	0,08
3	0,16	0,77	0,94	1	0,08	0,05

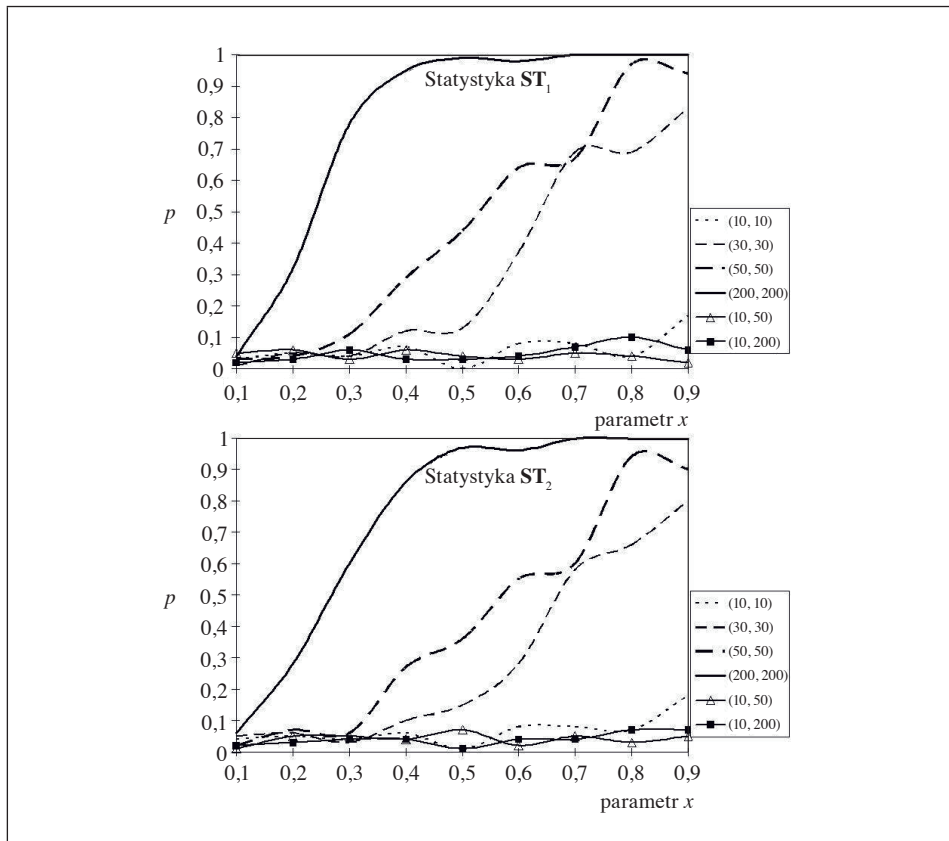
Źródło: obliczenia własne.



Rys. 7. Przypadek testowy 3

Źródło: obliczenia własne.

Przeprowadzone symulacje pokazują, że możliwa jest weryfikacja hipotezy H_0 za pomocą testu permutacyjnego z proponowanymi statystykami opartymi na różnicy wektorów własnych. Metoda ta jest bardziej uniwersalna niż znane testy parametryczne i nieparametryczne: pozwala na weryfikację bez konieczności badania jakichkolwiek założeń nie tylko w przypadkach, w których badane populacje, z których pobrano próby, różnią się wektorem wartości średnich czy macierzą kowariancji, ale także w przypadku innych odkształceń od rozkładu normalnego wielowymiarowego (asymetria, spłaszczenie). Pewne wątpliwości budzi praktyczna przydatność tej metody dla małych prób (o licznosci (10, 10) i w przypadkach dużej różnicy licznosci: (10, 50), (10, 200), gdzie prawdopodobieństwo odrzucenia jest bliskie 0 nawet przy znacznym wzroście parametru x .



Rys. 8. Przypadek testowy 5

Źródło: obliczenia własne.

Tabela 7. Wartości własne dla drugiej próby danych

Przypadek testowy	Wartości własne				
1	1,11	1,06	0,98	0,94	0,90
2, parametr $x = 0,9$	1,90	1,04	1,00	0,95	0,11
3, parametr $x = 0,9$	2,81	1,04	0,96	0,10	0,09
4, wariant 1	3,77	0,34	0,31	0,30	0,29
4, wariant 2	3,91	0,29	0,28	0,26	0,25
4, wariant 3	2,93	0,54	0,54	0,50	0,49
5, parametr $x = 0,9$	2,79	1,02	0,98	0,11	0,11

Źródło: obliczenia własne.

Ponieważ wyniki analiz wskazują na niższe prawdopodobieństwa dla statystyki ST_2 , przeprowadzono dodatkowe badania, śledząc możliwe przyczyny takiego zjawiska. W tabeli 7 przedstawiono wartości własne dla odkształcanych prób danych (drugie próby) analizowanych przypadków testowych. Prawdopodobnie gorsze zachowanie się testu dla statystyki ST_2 jest wynikiem pewnej ułomności symulowanych prób, których elipsoidy są bardzo regularne, z kilkoma osiami symetrii. W efekcie w wyjaśnianiu dominuje pierwszy wektor własny. Dalsze badania prowadzone na rzeczywistych pakietach danych zweryfikują zachowanie się obu statystyk testowych.

7. Podsumowanie

Rozwój możliwości obliczeniowych współczesnych komputerów zachęca do stosowania nowych metod statystycznych. Do takich metod niewątpliwie zalicza się testy permutacyjne. Artykuł prezentuje kolejną możliwość ich stosowania będącą alternatywą dla znanych metod porównywania dwóch populacji wielowymiarowych, niewymagającą badania założeń o rozkładzie ani homoskedastyczności. Zaletą tej metody jest możliwość weryfikacji za pomocą jednego testu hipotezy o braku różnic między badanymi próbami, niezależnie od kształtu elipsoid je reprezentujących. Metoda ta może być stosowana do porównywania kilku populacji albo śledzenia zmian w tej samej populacji. Nasuwa się pytanie, czy może ona służyć także do weryfikacji hipotezy o rozkładzie normalnym wielowymiarowym (badanej populacji) przez porównanie z inną populacją o rozkładzie normalnym. Próba odpowiedzi na to pytanie będzie przedmiotem dalszych badań.

Literatura

- Domański C., Pruska K. [2000], *Nieklasyczne metody statystyczne*, PWE, Warszawa.
- Good P.I. [2005], *Permutation Tests: A Practical Guide for Testing Hypotheses*, Springer-Verlag, New York.
- Hesterberg T. i in. [2003], *The Practice of Business Statistics, Companion Chapter 18 0 Bootstrap Methods and Permutation Tests*, H. Freeman, New York.
- Ito P.K. [1980], *Robustness of ANOVA and MANOVA Test Procedures [w:] Handbook of Statistics 1. Analysis of Variance*, red. P.R. Krishnaiah, North Holland, Amsterdam.

On the Comparison of Two Populations

A comparison of two populations is an interesting and very common statistical problem. It involves finding statistically significant differences based on given samples.

The most common way is to verify the hypothesis concerned the equality of certain, characteristic parameters, i.e. mean, standard deviation or fraction. The most efficient parametric tests need to fulfill assumptions about the normal distribution of examined populations. There are, however, cases where comparing “the shape” of multivariate populations could be crucial. Additionally, the distributions of tested populations are either unknown or cannot be treated as multivariate normal distributions.

This paper presents the results of investigations on the comparison of two populations where the differences between eigenvectors were implemented. Test statistics, based on the differences between first eigenvectors of tested populations, make it possible to examine the differences of a shape, regardless of its mean or standard deviation. They could be used, for example, to test the variability of a given phenomenon even with the trends. It was proposed to verify the hypotheses with permutation tests, where no assumptions about the distribution must be fulfilled. Doing so would make it possible to use different test statistics as well. At the end of the paper, the characteristics of the examined tests were estimated using Monte Carlo simulation.

Keywords: statistical test, permutation test, eigenvectors, Monte Carlo method.

Kamil Fijorek

Katedra Statystyki

Uniwersytet Ekonomiczny w Krakowie

Aproksymacja modelu regresji logistycznej Firtha za pomocą ważenia obserwacji

Streszczenie

W artykule przedstawiono model regresji logistycznej Firtha w kontekście wag przypisywanych przez metodę poszczególnym obserwacjom ze zbioru danych. Następnie dokonano przekształcenia modelu HLM do podobnej postaci. Na podstawie wniosków płynących z alternatywnego spojrzenia na model Firtha oraz HLM zaproponowano dwie metody aproksymacji modelu Firtha. Symulacyjnie zbadano jakość aproksymacji oraz omówiono praktyczne korzyści płynące z jej stosowania.

Słowa kluczowe: regresja logistyczna, redukcja obciążenia, całkowite rozdzielanie, estymacja przybliżona.

1. Wprowadzenie

Autor niniejszego artykułu porównał dwie metody estymacji modelu regresji logistycznej rozwiązujące problem całkowitego rozdzielania (*complete separation*) [Fijorek 2012]. Wspomniany problem występuje, gdy sukcesy i porażki mogą być całkowicie rozdzielone za pomocą jednej zmiennej niezależnej lub liniowej kombinacji kilku zmiennych niezależnych. W takim przypadku zastosowanie estymacji metodą największej wiarygodności (MNW) skutkuje uzyskaniem nieskończonych ocen parametrów modelu [Albert i Anderson 1984, s. 3–6; Firth 1993, s. 31–32;

Heinze 1999, s. 4–5]. Szansa pojawienia się całkowitego rozdzielenia jest tym większa, im mniejszy zbiór danych jest poddawany analizie.

Pierwszą metodą badaną w pracy [Fijorek 2012] jest propozycja G. Heinze i M. Schempera [2002] oparta na wynikach badań D. Firtha [1993]. Drugą propozycją jest metoda HLM (*hidden logistic model*) przedstawiona przez P.J. Rousseeuwa i A. Christmanna [2003] (w okresie późniejszym G. Tutz i F. Leitenstorfer [2006] opublikowali pracę o podobnym charakterze). Wnioski płynące z artykułu [Fijorek 2012] można podsumować w następujący sposób: oceny parametrów uzyskane metodą Firtha oraz HLM charakteryzują się mniejszym obciążeniem w porównaniu z ocenami MNW (w zakresie przeprowadzonych symulacji), przy czym metoda Firtha znacznie lepiej redukuje obciążenie niż HLM. Ponadto w modelu Firtha przedziały ufności wyznaczane za pomocą metody *profile likelihood* (PL) w porównaniu z przedziałami Walda uzyskują bliższy nominalnemu poziom pokrycia.

Praktyczna użyteczność modelu Firtha uzasadnia dalsze badania jego właściwości. Celem artykułu jest zaproponowanie metod aproksymacji modelu Firtha oraz wskazanie obszarów ich zastosowań w analizie danych empirycznych.

2. Model regresji logistycznej Firtha

W standardowym modelu regresji logistycznej przyjmuje się, że zmienna zależna $y_i \in \{0, 1\}$ ($i = 1, \dots, n$) podlega rozkładowi Bernoulliego z prawdopodobieństwem sukcesu $F(\mathbf{x}'_i \boldsymbol{\theta})$, gdzie $F(\cdot)$ jest dystrybuantą rozkładu logistycznego, x_i to p -wymiarowy wektor zmiennych objaśniających, a $\boldsymbol{\theta} \in \mathbb{R}^p$ to (zawierający wyraz wolny) p -wymiarowy wektor parametrów strukturalnych [Long 1997, s. 40–61].

W celu oszacowania parametrów modelu wyznacza się funkcję wiarygodności oraz jej logarytm:

$$L(\boldsymbol{\theta} | y_1, \dots, y_n) = \prod_{i=1}^n F(\mathbf{x}'_i \boldsymbol{\theta})^{y_i} [1 - F(\mathbf{x}'_i \boldsymbol{\theta})]^{1-y_i},$$

$$l(\boldsymbol{\theta} | y_1, \dots, y_n) = \sum_{i=1}^n y_i \ln F(\mathbf{x}'_i \boldsymbol{\theta}) + (1 - y_i) \ln [1 - F(\mathbf{x}'_i \boldsymbol{\theta})].$$

Następnie oblicza się pochodne cząstkowe logarytmu funkcji wiarygodności względem parametrów modelu:

$$s(\boldsymbol{\theta} | y_1, \dots, y_n) = \sum_{i=1}^n (y_i - F(\mathbf{x}'_i \boldsymbol{\theta})) \mathbf{x}_i.$$

Rozwiązanie układu równań $s(\boldsymbol{\theta} | y_1, \dots, y_n) = 0$ jest równoważne ze znalezieniem ocen parametrów maksymalizujących funkcję wiarygodności.

W przypadku modelu regresji logistycznej Firtha $s(\boldsymbol{\theta} | y_1, \dots, y_n)$ funkcję zastępuje się funkcją:

$$s^*(\boldsymbol{\theta} | y_1, \dots, y_n) = \sum_{i=1}^n \left(y_i - F(\mathbf{x}_i' \boldsymbol{\theta}) + h_i \left(\frac{1}{2} - F(\mathbf{x}_i' \boldsymbol{\theta}) \right) \right) \mathbf{x}_i,$$

gdzie h_i to diagonalne elementy macierzy $\mathbf{H} = \mathbf{W}^{\frac{1}{2}} \mathbf{X} (\mathbf{X}' \mathbf{W} \mathbf{X})^{-1} \mathbf{X}' \mathbf{W}^{\frac{1}{2}}$, \mathbf{X} to macierz danych, a \mathbf{W} jest macierzą diagonalną o wymiarach $n \times n$, której i -ty diagonalny element jest równy $F(\mathbf{x}_i' \boldsymbol{\theta})(1 - F(\mathbf{x}_i' \boldsymbol{\theta}))$. Modyfikacja postaci funkcji $s(\boldsymbol{\theta} | y_1, \dots, y_n)$ jest tożsama z modyfikacją funkcji wiarygodności:

$$L^*(\boldsymbol{\theta} | y_1, \dots, y_n) = L(\boldsymbol{\theta} | y_1, \dots, y_n) |\mathbf{I}_{\boldsymbol{\theta}}|^{\frac{1}{2}},$$

gdzie $\mathbf{I}_{\boldsymbol{\theta}}$ to macierz informacyjna postaci [Greene 2003, s. 670–673]:

$$\mathbf{I}_{\boldsymbol{\theta}} = - \sum_{i=1}^n \frac{\partial^2 l_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \sum_{i=1}^n F(\mathbf{x}_i' \boldsymbol{\theta})(1 - F(\mathbf{x}_i' \boldsymbol{\theta})) \mathbf{x}_i \mathbf{x}_i'.$$

$L^*(\boldsymbol{\theta} | y_1, \dots, y_n)$ bywa nazywana funkcją wiarygodności z karą (*penalized likelihood function*) [Heinze i Schemper 2002, Heinze 2006]. D. Firth [1993] udowodnił, że opisane powyżej podejście, poza rozwiązaniem problemu całkowitego rozdzielania, usuwa również obciążenie pierwszego rzędu towarzyszące metodzie największej wiarygodności.

3. Model regresji logistycznej z ukrytą zmienną objaśnianą (HLM)

W celu rozwiązania problemu całkowitego rozdzielania P.J. Rousseeuw i A. Christmann [2003] proponują natomiast założyć istnienie mechanizmu stochastycznego, który powoduje, że prawdziwe wartości zmiennej zależnej (y_i) są nieobserwowalne. W tym ujęciu przyjmuje się, że sukces jest obserwowany z prawdopodobieństwem δ_1 , co przekłada się na możliwość błędnego uznania sukcesu za porażkę z prawdopodobieństwem $1 - \delta_1$. Podobnie porażka jest obserwowana z prawdopodobieństwem $1 - \delta_0$, a jej błędna klasyfikacja występuje z prawdopodobieństwem δ_0 .

P.J. Rousseeuw i A. Christmann [2003, s. 317–318] pokazują, że estymacja parametrów modelu HLM sprowadza się do zastąpienia oryginalnych obserwacji pseudoobserwacjami \tilde{y}_i obliczonymi w następujący sposób:

$$\tilde{y}_i = (1 - y_i) \delta_0 + y_i \delta_1.$$

Następnie do tak utworzonych pseudoobserwacji można zastosować klasyczną metodę największej wiarygodności, której celem jest maksymalizacja funkcji wiarygodności względem wektora parametrów $\boldsymbol{\theta} \in \mathbb{R}^p$:

$$L(\boldsymbol{\theta} | \tilde{y}_1, \dots, \tilde{y}_n) = \prod_{i=1}^n F(\mathbf{x}_i' \boldsymbol{\theta})^{\tilde{y}_i} [1 - F(\mathbf{x}_i' \boldsymbol{\theta})]^{1 - \tilde{y}_i}.$$

Jak podają P.J. Rousseeuw i A. Christmann [2003, s. 320–323], precyzyjne wartości δ_0 oraz δ_1 mogą być poznane jedynie dla bardzo dużych zbiorów danych. W przypadku małych zbiorów danych można przyjąć, że $\delta_0 = 0,01$ i $\delta_1 = 0,99$. Jeśli założymy, że suma pseudoobserwacji powinna być równa sumie oryginalnych wartości zmiennej objaśnianej:

$$\hat{\pi} = \frac{1}{n} \sum_{i=1}^n \tilde{y}_i = \frac{1}{n} \sum_{i=1}^n y_i,$$

δ_0 oraz δ_1 przyjmują wartości:

$$\delta_0 = \frac{\tilde{\pi}\delta}{1+\delta} \quad \text{oraz} \quad \delta_1 = \frac{1+\tilde{\pi}\delta}{1+\delta}.$$

P.J. Rousseeuw i A. Christmann [2003, s. 322] zakładają, że $\delta = 0,01$ (w dalszej części pracy rezultaty oznaczone skrótem HLM wynikają z tego założenia). Należy zaznaczyć, że opisane sposoby doboru δ_0 oraz δ_1 nie mają silnego uzasadnienia teoretycznego. Można sądzić, że algorytm określania optymalnych wartości δ_0 oraz δ_1 powinien uwzględniać rozmiar próby oraz liczbę zmiennych objaśniających. Zaprezentowane w dalszej części artykułu wyniki badań dostarczają nowych informacji w tym zakresie.

4. Alternatywne spojrzenie na model Firtha oraz HLM

4.1. Propozycje aproksymacji modelu Firtha

Jak już wspomniano wcześniej, model regresji logistycznej Firtha może być zdefiniowany jako standardowy model regresji logistycznej ze zmodyfikowaną funkcją wiarygodności. Na model Firtha można też jednak spojrzeć z innej perspektywy i wykazać, że podejście to sprowadza się do zastąpienia każdej oryginalnej obserwacji y_i dwiema nowymi obserwacjami, tj. y_i oraz $1 - y_i$ z wagami wynoszącymi odpowiednio $1 + \frac{h_i}{2}$ oraz $\frac{h_i}{2}$ [Heinze i Schemper 2002, s. 2412]. W rezultacie otrzymuje się zbiór danych liczący $2n$ przypadków, w którym:

$$\mathbf{y}^* = (y_1, \dots, y_n, 1 - y_1, \dots, 1 - y_n),$$

$$\mathbf{X}^{*1} = (\mathbf{x}'_1, \dots, \mathbf{x}'_n, \mathbf{x}'_1, \dots, \mathbf{x}'_n),$$

a wagi poszczególnych przypadków to $\mathbf{h} = (1 + \frac{h_1}{2}, \dots, 1 + \frac{h_n}{2}, \frac{h_1}{2}, \dots, \frac{h_n}{2})$.

W kontekście modelu Firtha powyższa obserwacja nie ułatwia obliczeniowego aspektu procesu estymacji, gdyż nadal wagi są zależne od danych oraz –

co bardziej istotne – wagi zmieniają się wraz z kolejnymi iteracjami algorytmu optymalizującego funkcję wiarygodności. Korzystając z powyższego spostrzeżenia, można jednak zaproponować metodę aproksymacji modelu Firtha polegającą na wyznaczeniu stałych wag, które aproksymują element $\frac{h_i}{2}$. Jak wykazują D.W. Hosmer i S. Lemeshow [2000, s. 169]:

$$\sum_{i=1}^n h_i = p,$$

potencjalnym sposobem przybliżania h_i jest zatem uznanie, że $h_i \approx \frac{p}{n} = \bar{h}$.

Pierwsza propozycja aproksymacji modelu Firtha (aFirth) polega na zastosowaniu standardowej metody największej wiarygodności do zbioru danych $(\mathbf{y}^*, \mathbf{X}^*)$, w którym wagi pierwszych n obserwacji to $1 + \frac{\bar{h}}{2}$, a wagi kolejnych n obserwacji to $\frac{\bar{h}}{2}$. Konstrukcja wag powoduje, że rola sztucznych obserwacji maleje wraz ze wzrostem n , a przy założeniu stałego rozmiaru zbioru danych ich rola rośnie wraz ze wzrostem liczby zmiennych niezależnych. Ponadto należy podkreślić, że oryginalne obserwacje odgrywają dominującą rolę w procesie estymacji, a obserwacje sztuczne stosunkowo niewielką, np. jeśli $n = 100$ oraz $p = 5$, to waga sztucznej obserwacji wynosi 0,025, natomiast obserwacji oryginalnej 1,025.

Jak wynika z obliczeń autora (inspirowanych pracą [Tutz i Leitenstorfer 2006]), również model HLM, który w swej podstawowej formie zakłada transformację oryginalnych obserwacji do pseudoobserwacji, można zmodyfikować i zapisać w formie zbioru danych $(\mathbf{y}^*, \mathbf{X}^*)$ z wagami w_i :

$$\begin{cases} y_i = 1 \rightarrow w_i = \delta_1 \\ y_i = 0 \rightarrow w_i = 1 - \delta_0 \end{cases} \quad \text{dla } i = 1, \dots, n,$$

$$\begin{cases} 1 - y_i = 1 \rightarrow w_i = \delta_0 \\ 1 - y_i = 0 \rightarrow w_i = 1 - \delta_1 \end{cases} \quad \text{dla } i = n + 1, \dots, 2n.$$

Druga propozycja aproksymacji modelu Firtha polega na założeniu, że $\delta = \frac{\bar{h}}{2}$, oraz zastosowaniu standardowej metody największej wiarygodności do zbioru danych $(\mathbf{y}^*, \mathbf{X}^*)$ z wagami w_i . Propozycja ta będzie oznaczana jako zmodyfikowany HLM (mHLM).

4.2. Aspekt praktyczny aproksymacji modelu Firtha

Aby uzyskać oceny parametrów modelu regresji logistycznej za pomocą metody Firtha, trzeba posiadać specjalne oprogramowanie. Przykładem implementacji modelu Firtha jest biblioteka *logistf* [Heinze i Ploner 2004] dostępna w środowisku obliczeń statystycznych R [R Development Core Team... 2010]. Inne

pakiety statystyczne (z pewnymi wyjątkami) nie udostępniają modelu Firtha, lecz większość z nich dopuszcza ważenie obserwacji niecałkowitymi wagami. W tej sytuacji użytkownik może aproksymować model Firtha, stosując standardowy model regresji logistycznej (w środowisku R służy temu funkcja *glm*) z wagami obliczonymi na podstawie wyników niniejszego artykułu.

Istniejące implementacje modelu Firtha znajdują się we wczesnym stadium rozwoju i w rezultacie charakteryzują się stosunkowo ograniczoną funkcjonalnością. Przykładem może być wspomniana powyżej biblioteka *logistf*, która umożliwia estymację modelu oraz testowanie hipotez statystycznych, lecz już automatyczne metody doboru zmiennych nie są obsługiwane. Proponowana aproksymacja umożliwia podejście dwuetapowe, tzn. w pierwszym kroku użytkownik stosuje aproksymację modelu Firtha, uzyskując tym samym dostęp do funkcjonalności standardowych narzędzi estymacji regresji logistycznej, natomiast w drugim etapie użytkownik dokonuje estymacji finalnego modelu za pomocą procedury Firtha.

5. Badania symulacyjne oraz przykład empiryczny

5.1. Założenia badań symulacyjnych

W tej części opracowania za pomocą badań symulacyjnych następujące trzy metody estymacji modelu regresji logistycznej:

- metoda aproksymowanego Firtha (aFirth),
- zmodyfikowana metoda HLM (mHLM),
- standardowa metoda HLM,

zostaną porównane z wzorcem, którym w tym badaniu jest model Firtha. W porównaniach wykorzystano zbiory danych zawierające:

- wariant A – k ($k = 2, 4, 6, 8, 10, 12$) nieskorelowanych zmiennych losowanych z rozkładu $N(0, 1)$,
- wariant B – $\frac{k}{2}$ nieskorelowanych zmiennych losowanych z rozkładu $N(0, 1)$ oraz $\frac{k}{2}$ zmiennych zero-jedynkowych losowanych z rozkładu Bernoulliego (prawdopodobieństwo uzyskania 1 przyjęto na poziomie 0,5).

W toku symulacji generowano $R = 1000$ zestawów danych liczących 50, 75, 100 oraz 200 przypadków. Wartość y_i określano jako losową realizację z rozkładu Bernoulliego, w którym prawdopodobieństwo sukcesu jest równe $F(\mathbf{x}_i' \boldsymbol{\theta})$. W ramach badania symulacyjnego wektor parametrów modelu rozważono w dwóch wariantach:

- wariant 1 – $\theta_j = 1$ ($j = 0, \dots, k$),
- wariant 2 – $\theta_j = 0,5$ ($j = 0, \dots, k$),

gdzie θ_0 oznacza wyraz wolny. W rezultacie otrzymano cztery warianty symulacji oznaczone kolejno A1, A2, B1, B2.

Zachowanie się ocen parametrów określano, mierząc ich przeciętne względne obciążenie zgodnie ze wzorem:

$$B = \frac{1}{k} \sum_{j=0}^k \frac{\left[\frac{1}{R} \sum_{r=1}^R \hat{\theta}_j^r - \theta_j \right]}{\theta_j},$$

gdzie $\hat{\theta}_j^r$ to ocena j -tego parametru na podstawie r -tej symulacji.

5.2. Wyniki badań symulacyjnych

Wyniki przeprowadzonych badań symulacyjnych zestawiono w tabelach 1–4. Zgodnie z oczekiwaniami zaobserwowano, że ogólnie im mniejszy jest zbiór danych przy założeniu stałej liczby zmiennych objaśniających oraz im więcej zmiennych objaśniających przy założeniu stałego rozmiaru zbioru danych, tym większe jest obciążenie ocen parametrów. Ponadto – również zgodnie z oczekiwaniami – oceny parametrów modelu regresji logistycznej wyznaczone metodą Firtha charakteryzowały się bardzo niskim obciążeniem w przekroju niemal wszystkich wariantów symulacji. Niewielkie odstępstwa od braku obciążenia zaobserwowano w najbardziej ekstremalnych przypadkach, tzn. w przypadku $n = 50$, $n = 75$ oraz $k = 10$, $k = 12$. Wykonane symulacje pozwalają przewidywać, że korekta ocen parametrów modelu regresji logistycznej o obciążenie rzędu wyższego niż pierwszy nie będzie skutkować poprawą, którą można by uznać za istotną z praktycznego punktu widzenia. Interesujące rozważania teoretyczne na ten temat w klasie uogólnionych modeli liniowych (*generalized linear model*), których szczególnym przypadkiem jest model regresji logistycznej, przedstawili G. Cordeiro i L. Barroso [2007].

Model HLM w swej standardowej postaci w porównaniu z modelem Firtha wykazał bardzo słabą kontrolę obciążenia ocen parametrów. Dopiero przy $n = 200$ oraz niewielkiej liczbie zmiennych objaśniających obciążenie znajdowało się na akceptowalnym poziomie. Dużo lepszą kontrolą obciążenia wykazał się natomiast zmodyfikowany model HLM (mHLM), który w znacznej części symulacji dawał wyniki zbliżone do modelu Firtha. Interesujące jest, że propozycja aFirth nie dostarczyła zadowalających rezultatów. Aproksymacja ta, ogólnie rzecz biorąc, prowadziła do zbyt mocnej korekty obciążenia (oceny parametrów były mniejsze, niż powinny być), przy czym zjawisko było tym bardziej dotkliwe, im więcej zmiennych objaśniających znajdowało się w modelu (przy założeniu stałego rozmiaru zbioru danych).

Tabela 1. Wartości przeciętnego względnego obciążenia ocen parametrów – wariant A1 (w %)

Liczba obserwacji	Model	Liczba zmiennych (razem z wyrazem wolnym)					
		3	5	7	9	11	13
50	HLM	12,2	23,6	41,1	83,5	141,7	187,7
	mHLM	7,9	7,1	1,4	-4,6	-11,8	-21,1
	Firth	1,3	2,1	1,5	1,5	-0,4	-11,8
	aFirth	-2,6	-11,5	-22,6	-32,2	-40,2	-48,0
75	HLM	6,6	11,6	18,2	29,1	49,6	73,2
	mHLM	4,8	4,4	1,6	-2,5	-6,5	-13,0
	Firth	0,5	0,6	0,4	1,3	3,8	2,3
	aFirth	-1,6	-7,7	-15,7	-23,8	-30,7	-38,0
100	HLM	3,3	7,3	10,6	18,1	26,8	37,5
	mHLM	2,5	3,1	1,0	-0,3	-4,5	-8,8
	Firth	-0,7	0,1	-0,3	1,5	2,7	3,5
	aFirth	-2,1	-6,0	-12,2	-18,0	-25,3	-31,4
200	HLM	1,5	2,5	4,0	4,9	7,4	9,1
	mHLM	1,9	1,9	1,4	-0,2	-1,5	-3,9
	Firth	0,3	0,4	0,6	0,2	0,8	0,4
	aFirth	-0,4	-2,6	-5,5	-9,8	-13,8	-18,5

Źródło: opracowanie własne.

Tabela 2. Wartości przeciętnego względnego obciążenia ocen parametrów – wariant A2 (w %)

Liczba obserwacji	Model	Liczba zmiennych (razem z wyrazem wolnym)					
		3	5	7	9	11	13
50	HLM	6,2	14,5	24,7	39,7	66,8	118,5
	mHLM	4,0	7,0	8,8	9,5	7,8	5,6
	Firth	-2,3	-0,6	0,8	2,9	3,9	6,0
	aFirth	-2,1	-4,2	-7,5	-12,1	-18,3	-23,8
75	HLM	5,5	9,9	12,7	19,4	25,9	37,6
	mHLM	4,4	6,2	5,2	6,3	5,6	5,4
	Firth	0,2	0,9	-0,2	1,0	1,2	3,0
	aFirth	0,5	-0,9	-5,0	-7,8	-11,8	-15,7
100	HLM	2,9	6,8	9,0	12,5	16,5	21,2
	mHLM	2,4	4,6	4,4	4,5	4,3	3,6
	Firth	-0,7	0,6	0,2	0,4	0,6	0,8
	aFirth	-0,4	-0,5	-3,1	-5,8	-8,9	-12,3
200	HLM	0,6	2,8	4,0	4,9	6,4	6,9
	mHLM	0,8	2,5	2,8	2,5	2,6	1,4
	Firth	-0,7	0,5	0,6	0,3	0,5	-0,4
	aFirth	-0,5	0,1	-0,9	-2,5	-3,9	-6,5

Źródło: opracowanie własne.

Tabela 3. Wartości przeciętnego względnego obciążenia ocen parametrów – wariant B1 (w %)

Liczba obserwacji	Model	Liczba zmiennych (razem z wyrazem wolnym)					
		3	5	7	9	11	13
50	HLM	13,4	31,6	87,5	186,8	274,1	354,0
	mHLM	9,9	14,8	20,0	17,4	10,7	1,1
	Firth	0,3	0,7	1,6	-6,4	-19,6	-34,4
	aFirth	-4,0	-15,8	-29,2	-41,8	-52,0	-59,9
75	HLM	7,8	16,1	36,0	85,1	173,5	263,2
	mHLM	6,4	10,1	15,6	20,6	23,4	21,0
	Firth	0,7	0,8	2,1	1,9	-3,4	-16,4
	aFirth	-2,1	-10,5	-21,7	-33,5	-44,6	-53,8
100	HLM	5,2	10,3	18,6	38,7	86,2	181,1
	mHLM	4,6	7,2	10,1	14,7	20,3	25,8
	Firth	0,4	0,7	0,3	1,0	0,5	-2,7
	aFirth	-1,6	-7,7	-17,2	-28,0	-38,6	-47,7
200	HLM	1,4	3,3	7,0	10,1	16,6	27,2
	mHLM	1,7	2,8	5,3	6,2	9,0	12,4
	Firth	-0,3	-0,2	0,8	0,0	0,3	0,5
	aFirth	-1,2	-4,2	-8,7	-16,2	-24,4	-33,0

Źródło: opracowanie własne.

Tabela 4. Wartości przeciętnego względnego obciążenia ocen parametrów – wariant B2 (w %)

Liczba obserwacji	Model	Liczba zmiennych (razem z wyrazem wolnym)					
		3	5	7	9	11	13
50	HLM	9,8	18,0	29,1	71,1	163,6	232,0
	mHLM	7,9	11,9	14,8	22,3	28,2	29,4
	Firth	1,0	1,7	0,5	3,8	4,3	-2,1
	aFirth	0,7	-3,3	-10,8	-17,8	-26,5	-35,1
75	HLM	4,2	11,0	14,9	25,0	41,5	102,8
	mHLM	3,4	8,2	9,1	13,8	17,1	27,1
	Firth	-0,9	1,6	0,2	1,6	1,1	4,5
	aFirth	-1,0	-1,3	-6,7	-11,3	-18,9	-25,4
100	HLM	4,3	6,6	10,4	16,7	20,8	35,9
	mHLM	4,0	5,0	7,0	10,4	10,5	16,0
	Firth	0,7	0,2	0,4	1,7	-0,5	0,9
	aFirth	0,6	-1,7	-4,5	-8,0	-15,1	-20,9
200	HLM	1,5	2,6	3,5	6,1	8,9	10,6
	mHLM	1,7	2,3	2,6	4,4	6,1	6,6
	Firth	0,1	-0,1	-0,5	0,2	0,8	-0,3
	aFirth	0,1	-0,8	-2,7	-4,2	-6,6	-11,3

Źródło: opracowanie własne.

Podsumowując wyniki symulacji, należy stwierdzić, że model Firtha powinien być zawsze przedkładany nad inne rozważane w artykule podejścia. W sytuacji gdy jego zastosowanie nie jest możliwe, należy rekomendować stosowanie zmodyfikowanego modelu HLM (mHLM). W przypadku niewielkiej liczby zmiennych niezależnych różnica pomiędzy dwoma podejściami nie powinna być duża. Dla zbiorów danych liczących 200 i więcej obserwacji różnice pomiędzy podejściami przestają być zauważalne przy założeniu zrównoważonego stosunku liczby sukcesów do porażek.

5.3. Przykład empiryczny

W tabeli 5 zestawiono wyniki estymacji modelu regresji logistycznej za pomocą omówionych w artykule metod. Zbiór danych, który posłużył do estymacji modeli omawiają i analizują P.J. Rousseeuw i A. Christmann [2003, s. 327–328], a jego oryginalnym źródłem jest praca [Finney 1947]. Zmienną objaśnianą jest zmienna symbolizująca wystąpienie u badanej osoby wazokonstrykcji (termin oznacza zwężenie światła naczyń krwionośnych). Zmiennymi objaśniającymi są objętość powietrza wdychanego (*volume*) oraz tempo wdychania powietrza (*rate*).

Tabela 5. Wyniki estymacji modelu regresji logistycznej dla zbioru danych

Zmienna	MNW		HLM		mHLM		Firth		aFirth	
	$\hat{\theta}$	$SE(\hat{\theta})$	$\hat{\theta}$	$SE(\hat{\theta})$	$\hat{\theta}$	$SE(\hat{\theta})$	$\hat{\theta}$	$SE(\hat{\theta})$	$\hat{\theta}$	$SE(\hat{\theta})$
Wyraz wolny	-9,53	3,23	-9,13	3,09	-8,21	2,78	-7,83	2,66	-7,27	2,38
<i>Volume</i>	2,88	1,49	3,71	1,37	3,31	1,23	3,17	1,18	2,90	1,06
<i>Rate</i>	2,65	0,91	2,55	0,88	2,31	0,80	2,18	0,76	2,06	0,70

$SE(\hat{\theta})$ – błąd średni szacunku.

Źródło: opracowanie własne na podstawie: [Finney 1947].

Analizując wyniki estymacji, można stwierdzić, że do ocen metody Firtha najbardziej podobne są oceny zmodyfikowanej metody HLM, następnie aproksymowanej metody Firtha, dalej standardowej metody HLM, a najgorzej w tym porównaniu wypadają oceny metody największej wiarygodności.

6. Podsumowanie

W artykule przedstawiono model regresji logistycznej Firtha w kontekście wag przypisywanych przez metodę poszczególnym obserwacjom ze zbioru

danych oraz zaprezentowano przekształcenie modelu HLM do podobnej postaci. Na podstawie wniosków płynących z alternatywnego spojrzenia na model Firtha oraz HLM zaproponowano dwie metody aproksymacji modelu Firtha. Następnie symulacyjnie zbadano jakość aproksymacji oraz omówiono praktyczne korzyści płynące z jej stosowania.

Literatura

- Albert A., Anderson J.A. [1984], *On the Existence of Maximum Likelihood Estimates in Logistic Regression Models*, „Biometrika”, vol. 71.
- Cordeiro G., Barroso L. [2007], *A Third-order Bias Corrected Estimate in Generalized Linear Models*, „Test”, vol. 16, nr 1.
- Fijorek K. [2012], *Porównanie modeli regresji logistycznej odpornych na problem całkowitego rozdzielenia*, Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie, nr 884, Kraków.
- Finney D.J. [1947], *The Estimation from Individual Records of the Relationship between Dose and Quantal Response*, „Biometrika”, vol. 34.
- Firth D. [1993], *Bias Reduction of Maximum Likelihood Estimates*, „Biometrika”, vol. 80.
- Greene W.H. [2003], *Econometric Analysis*, Pearson Education, New Jersey.
- Heinze G. [1999], *The Application of Firth's Procedure to Cox and Logistic Regression*, Technical Report 10, Department of Medical Computer Sciences, Section of Clinical Biometrics, Vienna University, Vienna.
- Heinze G. [2006], *A Comparative Investigation of Methods for Logistic Regression with Separated or Nearly Separated Data*, „Statistics in Medicine”, vol. 25.
- Heinze G., Ploner M. [2004], *A SAS Macro, S-PLUS Library and R Package to Perform Logistic Regression without Convergence Problems*, Technical Report 2, Section of Clinical Biometrics, Department of Medical Computer Sciences, Medical University of Vienna, Vienna.
- Heinze G., Schemper M. [2002], *A Solution to the Problem of Separation in Logistic Regression*, „Statistics in Medicine”, vol. 21.
- Hosmer D.W., Lemeshow S. [2000], *Applied Logistic Regression*, John Wiley and Sons.
- Long J.S. [1997], *Regression Models for Categorical and Limited Dependent Variables*, Sage, Thousand Oaks.
- R Development Core Team. *R: A Language and Environment for Statistical Computing* [2010], R Foundation for Statistical Computing, Vienna.
- Rousseeuw P.J., Christmann A. [2003], *Robustness against Separation and Outliers in Logistic Regression*, „Computational Statistics and Data Analysis”, vol. 43.
- Tutz G., Leitenstorfer F. [2006], *Response Shrinkage Estimators in Binary Regression*, „Computational Statistics and Data Analysis”, vol. 50.

Firth's Logistic Regression Approximation by Weighting Observations

Firth's approach to a logistic regression is presented from the perspective of weighted data points. Hidden Logistic Model is reformulated accordingly and two approximations of Firth's procedure are introduced. A simulation study was conducted to investigate and compare the quality of the approximations.

Keywords: logistic regression, bias reduction, complete separation, approximate estimation.

Bogumiła Krzeszowska-Zakrzewska

Katedra Badań Operacyjnych
Uniwersytet Ekonomiczny w Katowicach

Zastosowanie programowania zero-jedynkowego w harmonogramowaniu czynności projektu

Streszczenie

Ze względu na znaczący wzrost zainteresowania zagadnieniami związanymi z zarządzaniem projektami coraz częściej w literaturze przedmiotu podejmowany jest problem harmonogramowania czynności projektu. Głównymi technikami harmonogramowania czynności projektu są: wykres Gantta, metoda PERT i metoda CPM. Techniki te dostarczają czytelnych harmonogramów optymalizujących czas trwania projektu, a także zapewniają kontrolę wykorzystania zasobów w projekcie. W praktyce harmonogramy projektów mogą być optymalizowane nie tylko ze względu na czas trwania projektu, lecz także ze względu na poziom wykorzystania zasobów czy przepływy pieniężne. Wspomniane wyżej techniki nie uwzględniają tych czynników w optymalizacji harmonogramu.

Celem niniejszego opracowania jest przedstawienie modeli matematycznych harmonogramowania czynności projektów. Przedstawione zostaną trzy modele: model optymalizujący czas trwania projektu, model optymalizujący poziom wykorzystania zasobów oraz model optymalizujący przepływy pieniężne w projekcie. Cechą wspólną prezentowanych modeli matematycznych jest binarna postać zmiennej. Dodatkowo każdy model posiada ograniczenia dotyczące relacji kolejnościowych wykonywania czynności oraz uwzględnia warunek, że każda czynność może zostać wykonana tylko raz.

Do rozwiązania przedstawionych modeli zastosowano programowanie zero-jedynkowe, które jest szczególnym przypadkiem programowania całkowitoliczbowego.

Słowa kluczowe: harmonogramowanie projektu, programowanie zero-jedynkowe, optymalizacja harmonogramów, optymalizacja wielokryterialna.

1. Wprowadzenie

Ze względu na rosnące znaczenie projektów we współczesnych organizacjach w ciągu ostatnich kilkudziesięciu lat nastąpił rozwój dziedziny zarządzania projektem. Według jednej z definicji projekt (często określany także jako przedsięwzięcie) to zbiór działań lub prac, które służą do realizacji celu jedyne­go i mierzalnego [Brandenburg 2002]. Każdy projekt składa się m.in. z takich elementów, jak: czynności (zadania bądź operacje, które trwają określony czas oraz wymagają zasobów do realizacji), zasoby (niezbędne do realizacji czynności) oraz relacje kolejnościowe (relacje odwzorowujące logiczną kolejność wykonywania czynności). Reasumując, projekt składa się z czynności, które generować mogą koszty lub przepływy pieniężne, są ograniczone przez relacje kolejnościowe, a do ich realizacji niezbędne są zasoby, których dostępność również jest ograniczona. Właśnie te składowe projektu sprawiają, że harmonogramowanie jest istotnym elementem procesu zarządzania projektem.

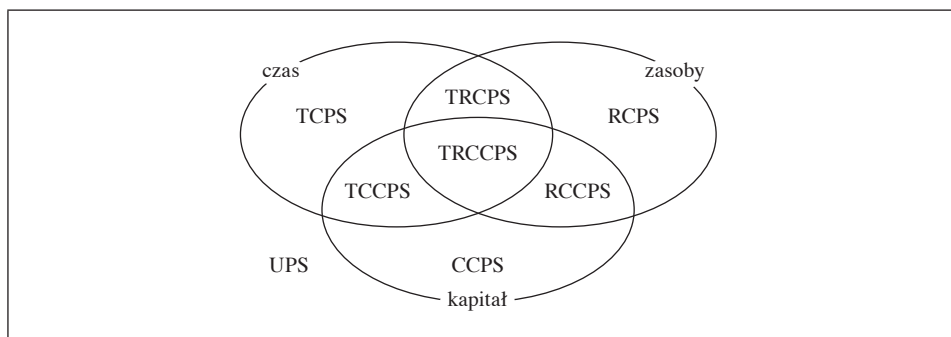
Harmonogram najogólniej określa kolejność i czas trwania poszczególnych czynności. Proces układania harmonogramów, czyli harmonogramowanie, można określić jako problem grupowania zasobów (lub zdarzeń) w pewnych punktach czasowych w danym okresie, w taki sposób aby osiągnąć określone cele i spełnić określone założenia [Norberciak 2002]. Podstawowymi technikami harmonogramowania projektu są metoda ścieżki krytycznej oraz metoda PERT [Brandenburg 2002]. Wykorzystując te techniki, można uzyskać czytelne harmonogramy (przez wykorzystanie wykresu Gantta) optymalizujące czas trwania projektu oraz zapewniające kontrolę wykorzystania zasobów. W praktyce optymalizować można nie tylko czas trwania projektów, lecz również wykorzystanie zasobów czy generowane przepływy pieniężne.

Celem niniejszego opracowania jest przedstawienie problemu harmonogramowania czynności projektu jako zdania programowania zero-jedynkowego. Programowanie zero-jedynkowe jest szczególnym przypadkiem programowania całkowitoliczbowego, w którym zmienne przyjmują wartości binarne.

2. Modele matematyczne harmonogramowania projektów

Szerokie zainteresowanie projektami spowodowało rozwój metod optymalizacyjnych w harmonogramowaniu projektów. Obecnie nie ma jednoznacznej klasyfikacji problemów harmonogramowania projektów. Wynika to ze zróżnicowania tych problemów. Próbę usystematyzowania wiedzy na temat modeli harmonogramowania projektów podjął A. Kosturbiac [2003], który przedstawił klasyfikację harmonogramów opartą na rodzajach ograniczeń (rys. 1) i kierunkach optymalizacji.

zacji (tabela 1). Zwrócił on uwagę na kluczową rolę ograniczeń w procesie planowania. Wyodrębnił problemy harmonogramowania: 1) bez ograniczeń, w którym kryterium optymalizacji stanowi czas trwania projektu lub przepływy pieniężne, 2) z ograniczeniami czasowymi, w którym kryterium optymalizacji stanowią zasoby lub przepływy pieniężne, 3) z ograniczeniami zasobowymi, 4) z ograniczonym kapitałem oraz 5) z wieloma ograniczeniami.



Oznaczenia: UPS (*unconstrained project scheduling*) – harmonogramowanie projektów bez ograniczeń, TCPS (*time constrained project scheduling*) – harmonogramowanie projektu z ograniczonym czasem, RCPS (*resource constrained project scheduling*) – harmonogramowanie projektu z ograniczonymi zasobami, CCPS (*capital constrained project scheduling*) – harmonogramowanie projektu z ograniczonym kapitałem, TRCPS (*time and resource constrained project scheduling*) – harmonogramowanie projektu z ograniczonym czasem i zasobami, TCCPS (*time and capital constrained project scheduling*) – harmonogramowanie projektu z ograniczonym czasem i kapitałem, RCCPS (*resource and capital constrained project scheduling*) – harmonogramowanie projektu z ograniczonymi zasobami i kapitałem, TRCCPS (*time, resource and capital constrained project scheduling*) – harmonogramowanie projektu z ograniczonym czasem, zasobami i kapitałem.

Rys. 1. Ograniczenia w modelach harmonogramowania projektów

Źródło: [Kostrubiec 2003].

Tabela 1. Kierunki optymalizacji w modelach harmonogramowania projektów

Kryterium oceny	UPS	TCPS	RCPS	CCPS	TRCPS	TCCPS	RCCPS	TRCCPS
Czasowe	v	v	v	v	v	v	v	v
Obciążenie zasobów	v	v	–	v	–	v	–	–
Koszt ^a	v	v	v	v	v	v	v	v
NPV	v	v	v	v	v	v	v	v

Oznaczenia: jak do rys. 1.

^a Dotyczy modeli ze zmiennym zapotrzebowaniem na zasoby i/lub kapitał.

Źródło: [Kostrubiec 2003].

Najogólniej ujmując, można stwierdzić, że ograniczenia oraz kierunki optymalizacji i oceny harmonogramu realizacji projektu wynikają z trzech podstawowych czynników: czasu, zasobów oraz kapitału. Każdy z modeli optymalizacyjnych harmonogramowania projektu posiada ponadto ograniczenia dotyczące relacji kolejnościowych oraz warunki brzegowe.

Najprostszym z modeli jednokryterialnych optymalizacji harmonogramów projektu jest model bez ograniczeń wynikających z czasu, zasobów lub kapitału. R. Russell [1970] przedstawił model, w którym maksymalizowana była wartość NPV. Założył, że w momencie rozpoczęcia czynności ponoszone są wydatki, natomiast zakończenie pewnej grupy czynności generuje wpływy.

Inne modele optymalizacji harmonogramów projektu to modele z jednym typem ograniczeń. Wyróżnić można trzy typy takich modeli: z ograniczeniami czasowymi, z ograniczeniami zasobowymi oraz z ograniczonym kapitałem [Kosturbiec 2003].

Pracę na temat harmonogramowania projektu z ograniczeniem czasowym przedstawili M. Vanhoucke, E. Demeulemeester i W. Herron [2002]. Rozważają oni problem z równomiernie rozłożonymi płatnościami. Celem przedstawionego w pracy modelu jest maksymalizacja wartości NPV. Ograniczeniem modelu jest czas realizacji projektu.

W grupie modeli harmonogramowania projektów z ograniczonymi zasobami można również wyróżnić modele bez przepływów pieniężnych (jak te wspomniane wcześniej), w których optymalizacji podlegają charakterystyki czasowe, oraz modele z przepływami pieniężnymi, w których optymalizacji podlegają charakterystyki ekonomiczne [Kosturbiec 2003].

Pracę na temat harmonogramowania projektów z ograniczonymi zasobami przedstawili np. M. Shouman i in. [2006]; zaprezentowali oni klasyczny model, w którym minimalizowany jest czas trwania projektu.

Problem optymalizacji harmonogramu projektu, w którym optymalizacji podlegały zdyskontowane przepływy pieniężne rozwiązali O. Icmei oraz S. Erenguc [1996].

T. Talbot [1982] przedstawił problem harmonogramowania projektu z ograniczonymi zasobami oraz z rozwiązaniami kompromisowymi dla zasobów i czasu (*time-resource tradeoff* – słowo *tradeoff* można rozumieć jako „coś za coś”; powiązania występujące w modelach harmonogramowania projektów mogą wywoływać pewne konflikty – poprawa jednego elementu może się wiązać z pogorszeniem innego, w tym przypadku: można wykonać projekt szybciej, ale wymaga to większego zaangażowania zasobów).

Ograniczenie kapitału w harmonogramowaniu projektu wprowadzili do modelu R. Doersch i J. Patterson [1977]. Założyli oni w modelu ograniczoną dostępność kapitału w momencie rozpoczęcia projektu. Wielkość dostępnego kapitału zmienia się w czasie w zależności od przepływów generowanych przez czynności, które zostały umieszczone w harmonogramie.

W literaturze przedmiotu spotkać można również opracowania na temat modeli harmonogramowania projektów z wieloma ograniczeniami. Modele te to najczęściej kombinacje jednokryterialnych modeli harmonogramowania projektów.

Pracę, w której model zawiera wiele ograniczeń, zaprezentowali M. Vanhoucke, E. Demeulemeester i W. Herroelen [2001]. Przedstawili oni problem maksymalizacji NPV projektu przy ograniczonych zasobach oraz ograniczonym czasie.

Problem z ograniczonymi zasobami oraz czasem przedstawili w swojej pracy M. Bartusch, R. Morhring oraz F. Readermacher [1988]. W zaprezentowanym modelu minimalizowany jest wektor składający się z czasów zakończenia poszczególnych czynności. Autorzy zaproponowali dwa typy ograniczeń – czasowe oraz zasobowe. W przypadku ograniczeń czasowych zakłada się, że czynność rozpocznie się w tzw. oknie czasowym (czas pomiędzy najwcześniejszym i najpóźniejszym możliwym rozpoczęciem czynności).

L. Bianco, P. Dell’Olmo i M. Speranza [1998] skonstruowali model z ograniczonymi zasobami oraz kapitałem. Każda czynność w projekcie może być realizowana na różne sposoby, z czym związane są określone koszty realizacji. Budżet projektu stanowi ograniczenie wyboru sposobu realizacji czynności. Dodatkowo wprowadzono ograniczenie, że czynności wykorzystujące ten sam zasób nie mogą być realizowane jednocześnie. Funkcja celu w modelu dotyczy minimalizacji czasu trwania projektu.

Powyżej przedstawione zostały modele optymalizacyjne harmonogramowania projektów, które zostały sklasyfikowane ze względu na liczbę oraz rodzaj ograniczeń. Nie ma jak dotąd opracowań, które podejmowałyby problem harmonogramowania projektu i zawierałyby wszystkie typy ograniczeń: zasobów, czasu i kosztów. Teoretycznie możliwe jest zbudowanie i rozwiązanie takiego modelu.

Powszechnie stosowane modele harmonogramowania projektów są modelami jednokryterialnymi. Niewiele opracowań dotyczących tej problematyki podejmuje problem wielokryterialności modeli optymalizujących harmonogramy.

3. Programowanie zero-jedynkowe w harmonogramowaniu projektów

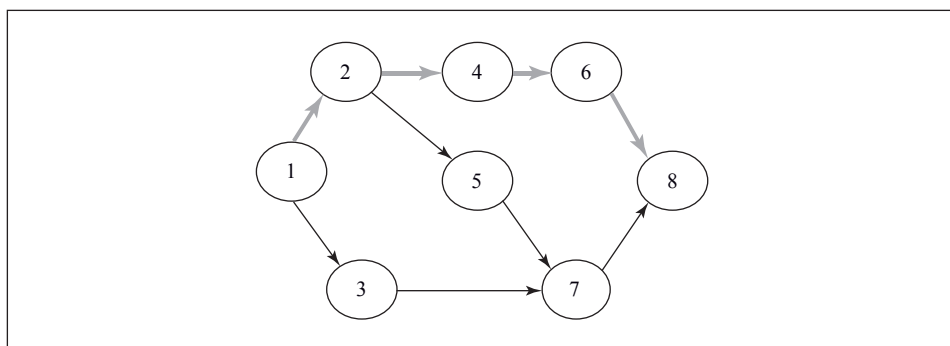
3.1. Opis problemu harmonogramowania projektu

Problem harmonogramowania czynności projektu przedstawiony w niniejszym opracowaniu można opisać w następujący sposób: należy utworzyć harmonogram projektu przez wyznaczenie czasów rozpoczęcia oraz zakończenia poszczególnych czynności, dysponując określonymi środkami przy ograniczeniach dotyczących relacji kolejnościowych oraz poziomu dostępności zasobów odnawialnych w poszczególnych jednostkach czasu.

W dalszych rozważaniach posłużono się następującym przykładem.

Przykład

Dany jest projekt składający się z 9 czynności. Projekt przedstawiony został w postaci sieci zdarzeń *Activity on Arc* (AOA) (rys. 2). W tabeli 2 zaprezentowano czas trwania poszczególnych czynności, wymagane zasoby odnawialne (przyjęto istnienie jednego takiego zasobu) oraz przepływy pieniężne netto generowane przez czynności w poszczególnych okresach ich trwania.



Rys. 2. Sieć czynności dla przykładu

Źródło: opracowanie własne.

Tabela 2. Dane do przykładu

Czynność	Czas trwania	ES	EF	LS	LF	Zapas	Czynność krytyczna	Zasoby odnawialne	Przepływy pieniężne netto w jednostce czasu			
									t = 1	t = 2	t = 3	t = 4
1–2	2	0	2	0	2	0	tak	2	-4	-4	0	0
1–3	4	0	4	1	5	1	nie	1	-3	-3	-2	-1
2–4	1	2	3	2	3	0	tak	2	-1	0	0	0
2–5	2	2	4	5	7	3	nie	3	1	1	0	0
3–7	3	4	7	5	8	1	nie	4	3	3	4	0
5–7	1	4	5	7	8	3	nie	2	5	5	6	0
4–6	4	3	7	3	7	0	tak	1	7	7	8	8
6–8	3	7	10	7	10	0	tak	3	8	9	8	0
7–8	2	7	9	8	10	1	nie	2	10	12	0	0

Oznaczenia: ES (*early start*) – najwcześniejszy moment rozpoczęcia czynności, EF (*early finish*) – najwcześniejszy moment zakończenia czynności, LS (*late start*) – najpóźniejszy moment rozpoczęcia czynności, LF (*late finish*) – najpóźniejszy moment zakończenia czynności.

Źródło: opracowanie własne.

Dla takiej sieci projektu metodą ścieżki krytycznej wyznaczone zostały najwcześniejszy oraz najpóźniejszy czas rozpoczęcia i zakończenia czynności projektu. Na tej podstawie wyznaczona została ścieżka krytyczna (rys. 3). Zgodnie z uzyskanymi wynikami projekt powinien zakończyć się po 10 jednostkach czasu. Wyniki uzyskane metodą ścieżki krytycznej porównane zostaną z wynikami uzyskanymi przez zastosowanie matematycznych modeli harmonogramowania projektów.

Dodatkowo w przykładzie przyjęto założenie, że w jednostce czasu dysponujemy 5 jednostkami zasobu odnawialnego, a czas na realizację przedsięwzięcia został ograniczony do 15 jednostek czasu.

Do budowy matematycznych modeli harmonogramowania projektów przyjęte zostały następujące założenia ogólne:

- projekt składa się z $j = 1, \dots, J$ czynności,
- czas realizacji przedsięwzięcia jest ograniczony do T ,
- $t = 0, \dots, T$ – zbiór okresów, gdzie T – maksymalny możliwy czas realizacji projektu,
- czynności projektu umieszczone zostały w postaci sieci czynności typu AOA,
- F_j – czas zakończenia czynności j projektu,
- relacje kolejnościowe typu zakończenie–początek (S_{ij} – zbiór bezpośrednich poprzedników i czynności j),
- F_{ij} – czas zakończenia poprzednika i czynności j ,
- $x_{ij} = \{0, 1\}$, $x_{ij} = 1$, gdy czynność j projektu kończy się w czasie t ,
- d_j – czas trwania czynności j ,
- $k = 1, \dots, K$ – zbiór zasobów odnawialnych,
- r_{jk} – wielkość zasobu odnawialnego k wymaganego przez czynność j .

3.2. Minimalizacja czasu trwania projektu jako zadanie programowania zero-jedynkowego

Pierwszy rozpatrywany model dotyczy minimalizacji czasu trwania projektu. Przedstawione zostaną dwa modele dla tego problemu – powszechnie stosowany w literaturze oraz propozycja własna modelu mającego na celu optymalizację czasu trwania projektu.

Pierwszy z modeli można przedstawić następująco:

$$\sum_{j=1}^J \sum_{t=1}^T t \times x_{jt} \rightarrow \min \quad (j = 1, \dots, J, t = 1, \dots, T), \quad (1)$$

przy ograniczeniach:

$$F_j - d_j \geq F_{ij} \quad (j = 1, \dots, J, t = 1, \dots, T, i \in S_{ij}), \quad (2)$$

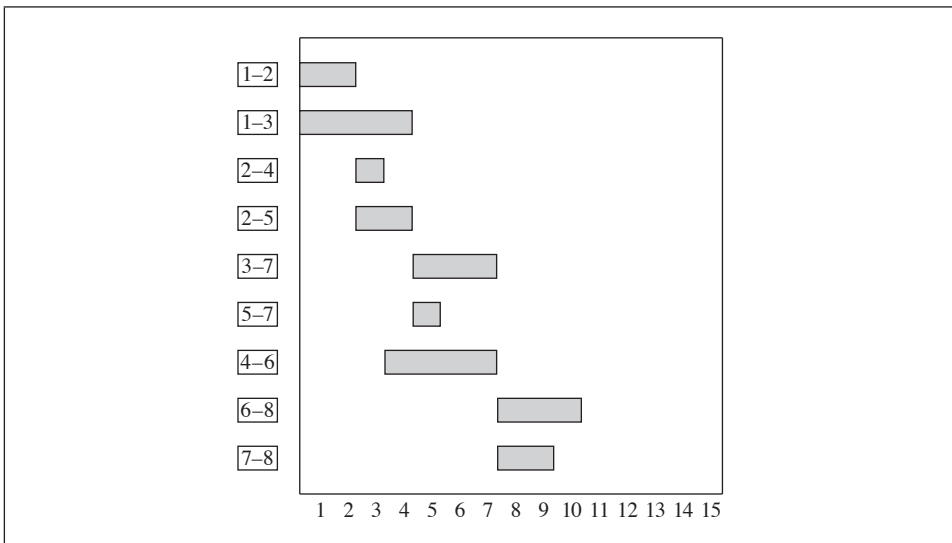
$$\sum_{j=1}^J r_{jk} \cdot x_{jt} \leq R_{kt} \quad (j = 1, \dots, J, t = 1, \dots, T, k = 1, \dots, K), \quad (3)$$

$$\sum_{t=1}^T x_{jt} = 1 \quad (j = 1, \dots, J, t = 1, \dots, T), \quad (4)$$

$$x_{jt} = \{0, 1\} \quad (j = 1, \dots, J, t = 1, \dots, T). \quad (5)$$

W przedstawionym modelu funkcja celu (1) polega na minimalizacji sumy iloczynu jednostki czasu i zmiennej decyzyjnej [Talbot 1982]. Ponieważ zmienna decyzyjna przyjmuje wartość 1 dla jednostki czasu, w której następuje jej zakończenie, można stwierdzić, że funkcja celu to po prostu suma czasów zakończeń poszczególnych czynności. W modelu dodano: ograniczenie dotyczące relacji kolejnościowych (2), które w rozpatrywanym przykładzie są relacjami typu zakończenie–początek, ograniczenie dotyczące wykorzystania zasobów odnawialnych w poszczególnych jednostkach czasu (3) oraz ograniczenie dotyczące zmiennej decyzyjnej (4), które polega na tym, że każda czynność może zakończyć się tylko raz, a więc suma zmiennych po jednostkach czasu dla każdej czynności powinna wynosić 1. Zmienna decyzyjna przyjmuje wartości 1 lub 0 (5).

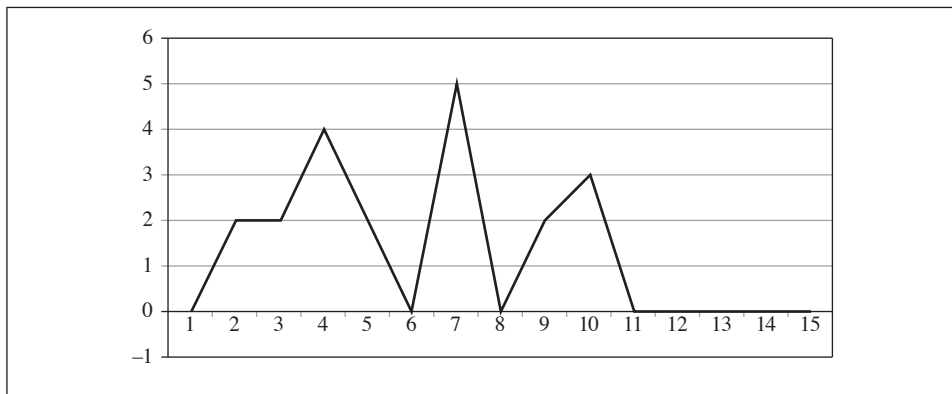
Przedstawiony powyżej model został wykorzystany do uzyskania harmonogramu dla rozpatrywanego przykładu. Rozwiązując powyższe zadanie, uzyskano następujące wyniki. Wartość funkcji celu wyniosła $z = 51$, co oznacza, że suma czasów zakończeń poszczególnych czynności projektu wynosi 51. Niestety wartość ta nie obrazuje, w jakim czasie zakończony zostanie projekt i czy zostanie on zakończony zgodnie z czasem, jaki uzyskaliśmy przez zastosowanie metody ścieżki krytycznej. Informację taką możemy uzyskać, analizując uzyskany harmonogram projektu (rys. 3).



Rys. 3. Minimalizacja czasu trwania projektu – harmonogram

Źródło: opracowanie własne.

Analiza uzyskanego harmonogramu pozwoliła na określenie momentu zakończenia projektu, który wynosi 10. Oznacza to, że harmonogram zakłada terminową realizację projektu. Rys. 4 przedstawia wykorzystanie zasobu odnawialnego w poszczególnych jednostkach czasu.



Rys. 4. Minimalizacja czasu trwania projektu – poziom wykorzystania zasobów

Źródło: opracowanie własne.

Ze względu na wskazaną wadę funkcji celu zaprezentowanego modelu zaproponowano optymalizację czasu trwania projektu jako zadania minimalizacji opóźnień projektu, co zostało zapisane w postaci następującego modelu:

$$\sum_{t=1}^T \sum_{j=1}^J \max\{0, t - LF_j\} \cdot x_{jt} \rightarrow \min \quad (j = 1, \dots, J, t = 1, \dots, T), \quad (6)$$

przy ograniczeniach:

$$F_j - d_j \geq F_{ij} \quad (j = 1, \dots, J, t = 1, \dots, T, i \in S_{ij}), \quad (7)$$

$$\sum_{j=1}^J r_{jk} \cdot x_{jt} \leq R_{kt} \quad (j = 1, \dots, J, t = 1, \dots, T, k = 1, \dots, K), \quad (8)$$

$$\sum_{t=1}^T x_{jt} = 1 \quad (j = 1, \dots, J, t = 1, \dots, T), \quad (9)$$

$$x_{jt} \in \{0, 1\} \quad (j = 1, \dots, J, t = 1, \dots, T). \quad (10)$$

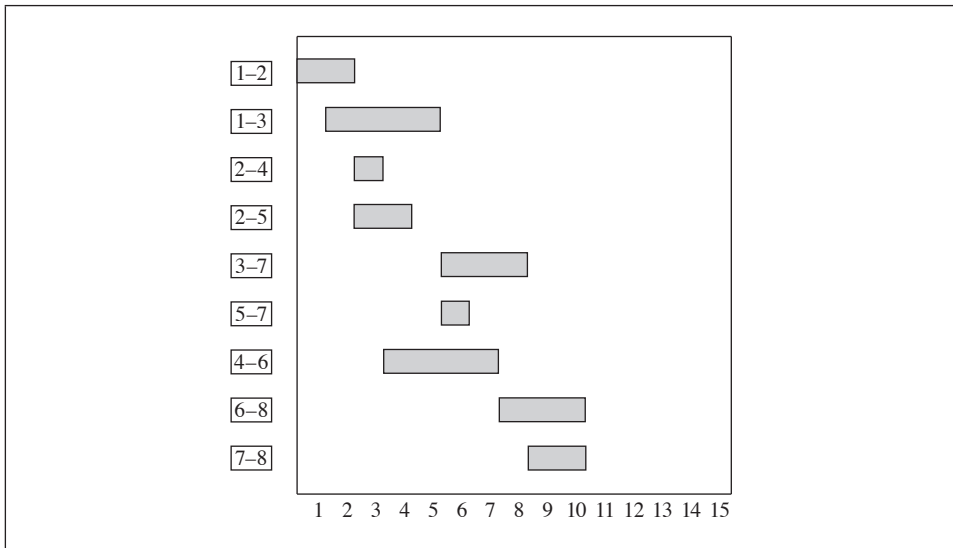
Funkcja celu (6) polega na minimalizacji opóźnień czynności projektu. Opóźnienie zostało określone jako różnica pomiędzy faktycznym czasem zakończenia czynności a najpóźniejszym czasem zakończenia czynności uzyskanym przez zastosowanie metody ścieżki krytycznej. Model ma ograniczenia (7)–(10), jak model poprzedni.

Po podstawieniu danych z przykładu funkcja celu ma następującą postać:

$$\begin{aligned}
 & x_{13}+2x_{14}+x_{34}+3x_{15}+2x_{35}+4x_{16}+x_{26}+3x_{36}+5x_{17}+2x_{27}+4x_{37}+ \\
 & 6x_{18}+3x_{28}+5x_{38}+x_{48}+x_{78}+7x_{19}+4x_{29}+6x_{39}+2x_{49}+x_{59}+x_{69}+2x_{79}+ \\
 & 8x_{110}+5x_{210}+7x_{310}+3x_{410}+2x_{510}+2x_{610}+3x_{710}+ \\
 & 9x_{111}+6x_{211}+8x_{311}+4x_{411}+3x_{511}+3x_{611}+4x_{711}+x_{811}+x_{911}+ \\
 & 10x_{112}+7x_{212}+9x_{312}+5x_{412}+4x_{512}+4x_{612}+5x_{712}+2x_{812}+2x_{912}+ \\
 & 11x_{113}+8x_{213}+10x_{313}+6x_{413}+5x_{513}+5x_{613}+6x_{713}+3x_{813}+3x_{913}+ \\
 & 12x_{114}+9x_{214}+11x_{314}+7x_{414}+6x_{514}+6x_{614}+7x_{714}+4x_{814}+4x_{914}+ \\
 & 13x_{115}+10x_{215}+12x_{315}+8x_{415}+7x_{515}+7x_{615}+8x_{715}+5x_{815}+5x_{915} \rightarrow \min
 \end{aligned}$$

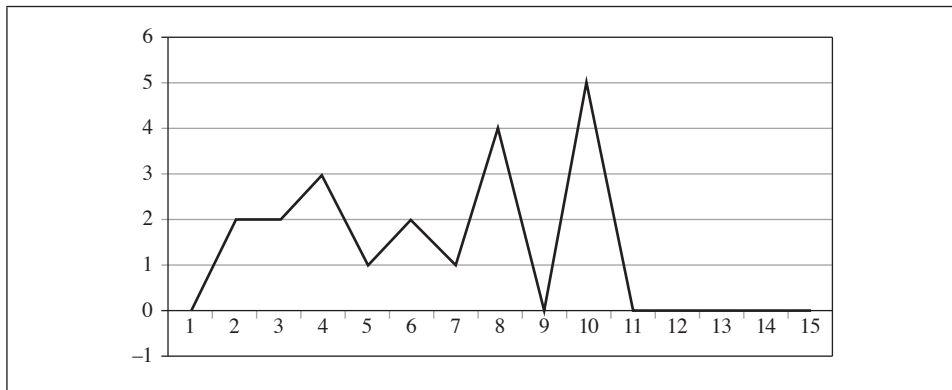
Analizując powyższą postać funkcji celu, można wywnioskować, że jeżeli zmienna x_{13} przyjmie wartość 1, co będzie oznaczało, że czynność pierwsza zakończy się w trzeciej jednostce czasu, odnotujemy opóźnienie na poziomie jednej jednostki czasu. Analogicznie jeżeli zmienna x_{14} przyjmie wartość 1, co będzie oznaczało, że czynność pierwsza zakończy się w czwartej jednostce czasu, wówczas odnotujemy opóźnienie na poziomie dwóch jednostek czasu.

Przedstawiony model optymalizujący czas trwania projektu przez minimalizację opóźnień projektu został wykorzystany do uzyskania harmonogramu dla omawianego przykładu.



Rys. 5. Minimalizacja opóźnień projektu – harmonogram

Źródło: opracowanie własne.



Rys. 6. Minimalizacja opóźnień projektu – poziom wykorzystania zasobów

Źródło: opracowanie własne.

Rozwiązując powyższe zadanie, uzyskano następujące wyniki. Wartość funkcji celu wyniosła $z = 0$. Taka wartość funkcji celu bezpośrednio informuje, że projekt realizowany jest terminowo w stosunku do czasu uzyskanego przy zastosowaniu metody ścieżki krytycznej, co można potwierdzić, analizując uzyskany harmonogram (rys. 5). W przypadku zastosowania funkcji kryterium, mającej na celu minimalizację sumy czasów zakończeń, poszczególne czynności harmonogramowane są zgodnie z najwcześniejszymi czasami rozpoczęcia i zakończenia uzyskanymi przez zastosowanie metody ścieżki krytycznej, podczas gdy zastosowanie zaproponowanej postaci funkcji celu prowadzi do umieszczenia czynności w harmonogramie zgodnie z najpóźniejszymi czasami rozpoczęcia i zakończenia ze ścieżki krytycznej. Rys. 6 przedstawia wykorzystanie zasobu odnawialnego w poszczególnych jednostkach czasu.

3.3. Minimalizacja poziomu wykorzystania zasobów odnawialnych jako zadanie programowania zero-jedynkowego

Kolejne zadanie jest związane z problemem optymalizacji poziomu wykorzystania zasobów. Problem ten nie jest często podejmowany w literaturze, a jego rozwiązanie polega zazwyczaj na minimalizacji odchylenia pomiędzy poziomem zasobów dostępnych a poziomem zasobów wykorzystanych [Viana i de Sousa 2000].

W niniejszym opracowaniu optymalizacja wykorzystania zasobów w projekcie polega na minimalizacji poziomu wykorzystania zasobów w poszczególnych jednostkach czasu, a problem ten zapisany może być w następujący sposób:

$$\max_{t=1, \dots, T} \left[\sum_{j=1}^J r_{jk} \cdot x_{jt} \right] \rightarrow \min (j = 1, \dots, J, t = 1, \dots, T, k = 1, \dots, K), \quad (11)$$

przy ograniczeniach:

$$F_j - d_j \geq F_{ij} \quad (j = 1, \dots, J, t = 1, \dots, T, i \in S_{ij}), \quad (12)$$

$$\sum_{t=1}^T x_{jt} = 1 \quad (j = 1, \dots, J, t = 1, \dots, T), \quad (13)$$

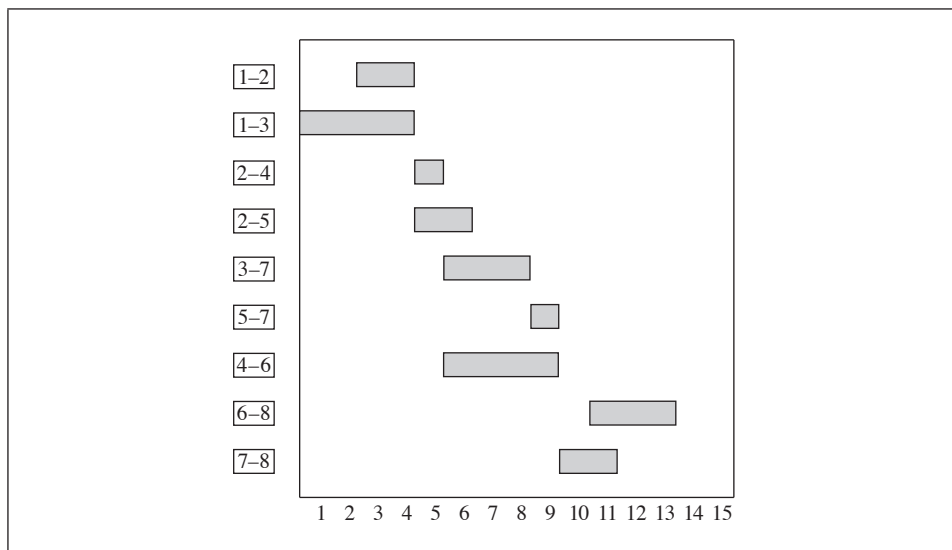
$$x_{jt} = \{0, 1\} \quad (j = 1, \dots, J, t = 1, \dots, T). \quad (14)$$

Funkcja kryterium (11) ma na celu minimalizację poziomu wykorzystania zasobów odnawialnych w poszczególnych jednostkach czasu. Ograniczenia (12)–(14) są takie same jak w poprzednich modelach. Ograniczenie dotyczące dostępności zasobów odnawialnych w poszczególnych jednostkach czasu w modelach optymalizujących poziom wykorzystania zasobów jest pomijane.

$$\begin{aligned} & \text{Max}\{(2x_{11}+x_{21}+2x_{31}+3x_{41}+4x_{51}+2x_{61}+x_{71}+3x_{81}+2x_{91}), \\ & (2x_{12}+x_{22}+2x_{32}+3x_{42}+4x_{52}+2x_{62}+x_{72}+3x_{82}+2x_{92}), \\ & (2x_{13}+x_{23}+2x_{33}+3x_{43}+4x_{53}+2x_{63}+x_{73}+3x_{83}+2x_{93}), \\ & (2x_{14}+x_{24}+2x_{34}+3x_{44}+4x_{54}+2x_{64}+x_{74}+3x_{84}+2x_{94}), \\ & (2x_{15}+x_{25}+2x_{35}+3x_{45}+4x_{55}+2x_{65}+x_{75}+3x_{85}+2x_{95}), \\ & (2x_{16}+x_{26}+2x_{36}+3x_{46}+4x_{56}+2x_{66}+x_{76}+3x_{86}+2x_{96}), \\ & (2x_{17}+x_{27}+2x_{37}+3x_{47}+4x_{57}+2x_{67}+x_{77}+3x_{87}+2x_{97}), \\ & (2x_{18}+x_{28}+2x_{38}+3x_{48}+4x_{58}+2x_{68}+x_{78}+3x_{88}+2x_{98}), \\ & (2x_{19}+x_{29}+2x_{39}+3x_{49}+4x_{59}+2x_{69}+x_{79}+3x_{89}+2x_{99}), \\ & (2x_{110}+x_{210}+2x_{310}+3x_{410}+4x_{510}+2x_{610}+x_{710}+3x_{810}+2x_{910}), \\ & (2x_{111}+x_{211}+2x_{311}+3x_{411}+4x_{511}+2x_{611}+x_{711}+3x_{811}+2x_{911}), \\ & (2x_{112}+x_{212}+2x_{312}+3x_{412}+4x_{512}+2x_{612}+x_{712}+3x_{812}+2x_{912}), \\ & (2x_{113}+x_{213}+2x_{313}+3x_{413}+4x_{513}+2x_{613}+x_{713}+3x_{813}+2x_{913}), \\ & (2x_{114}+x_{214}+2x_{314}+3x_{414}+4x_{514}+2x_{614}+x_{714}+3x_{814}+2x_{914}), \\ & (2x_{115}+x_{215}+2x_{315}+3x_{415}+4x_{515}+2x_{615}+x_{715}+3x_{815}+2x_{915})\} \rightarrow \min \end{aligned}$$

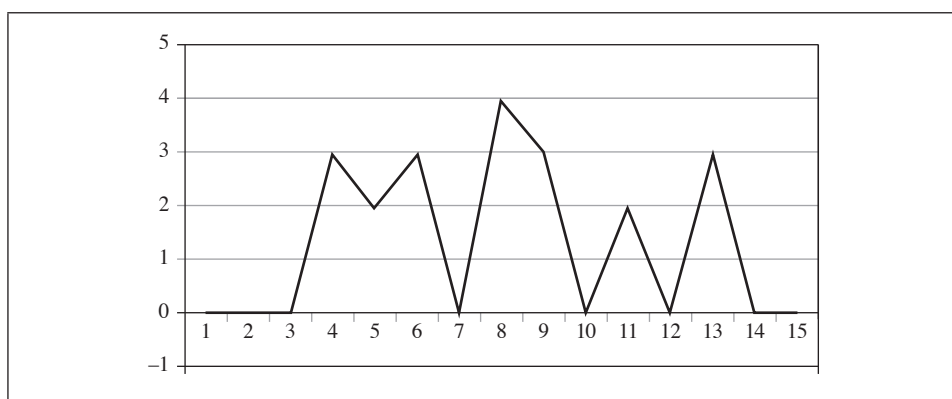
Rozwiązując powyższe zadanie, uzyskano następujące wyniki. Wartość funkcji celu wyniosła $z = 4$. Oznacza to, że maksymalny poziom wykorzystania zasobów odnawialnych w jednostce czasu wynosi 4, co potwierdza rys. 8. Analizując uzyskany harmonogram (rys. 7), możemy zauważyć, że obniżenie poziomu wykorzystania zasobów z 5 (takie ograniczenie nałożone było na poprzednie modele) do 4, jakie uzyskano przez rozwiązanie przedstawionego modelu matematycznego, pociąga za sobą wydłużenie czasu trwania projektu.

W aktualnym harmonogramie projekt kończy się po 13 jednostkach czasu, co oznacza, że jest on opóźniony o 3 jednostki czasu w stosunku do czasu zakończenia projektu, jaki uzyskano za pomocą metody ścieżki krytycznej.



Rys. 7. Minimalizacja poziomu wykorzystania zasobów – harmonogram

Źródło: opracowanie własne.



Rys. 8. Minimalizacja poziomu wykorzystania zasobów – poziom wykorzystania zasobów

Źródło: opracowanie własne.

3.4. Maksymalizacja zdyskontowanych przepływów pieniężnych w projekcie jako zadanie programowania zero-jedynkowego

Kierunki optymalizacji w modelach harmonogramowania projektów oprócz czasu i zasobów mogą dotyczyć również przepływów pieniężnych generowanych przez czynności projektu. Problem optymalizacji przepływów pieniężnych w problemie harmonogramowania projektu może być przedstawiony w postaci następującego modelu [Icmeli i Erenguc 1996]:

$$\sum_{i=1}^J \left[\sum_{t=1}^{d_j} [cf_{jt} \cdot e^{\alpha(d_j-t)}] \cdot e^{-\alpha F_j} \right] \rightarrow \max \quad (j = 1, \dots, J, t = 1, \dots, T), \quad (15)$$

przy ograniczeniach:

$$F_j - d_j \geq F_{ij} \quad (j = 1, \dots, J, t = 1, \dots, T, i \in S_{ij}), \quad (16)$$

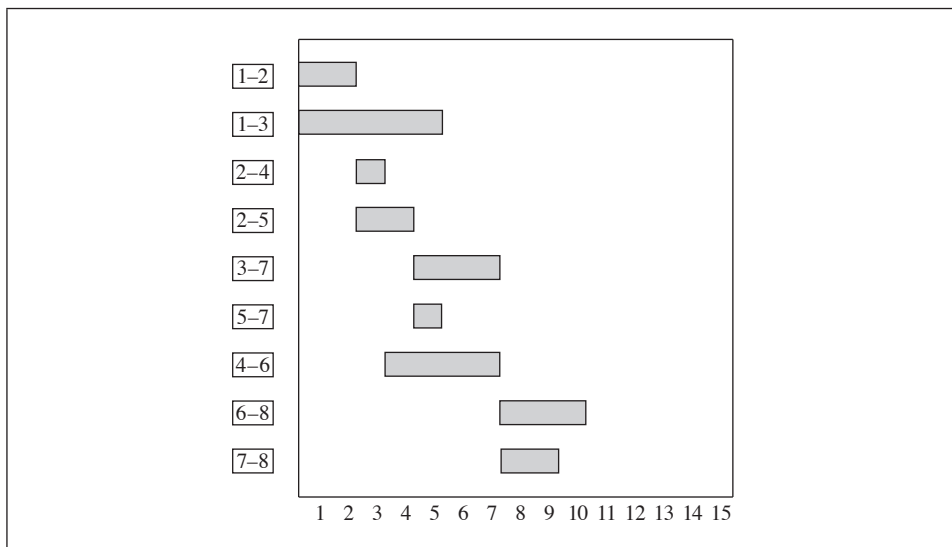
$$\sum_{j=1}^J r_{jk} \cdot x_{jt} \leq R_{kt} \quad (j = 1, \dots, J, t = 1, \dots, T, k = 1, \dots, K), \quad (17)$$

$$\sum_{t=1}^T x_{jt} = 1 \quad (j = 1, \dots, J, t = 1, \dots, T), \quad (18)$$

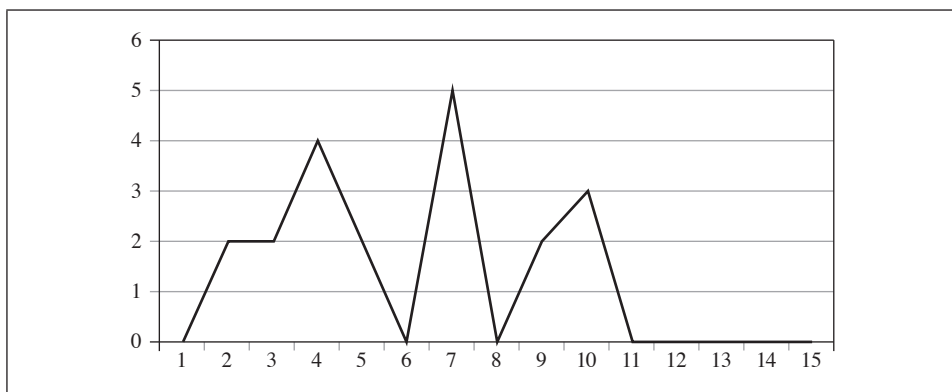
$$x_{jt} = \{0, 1\} \quad (j = 1, \dots, J, t = 1, \dots, T). \quad (19)$$

W funkcji celu (15) maksymalizowane są zdyskontowane przepływy pieniężne generowane przez poszczególne czynności projektu. Przepływy te są zależne od czasu trwania czynności oraz momentu jej zakończenia. Ograniczenia (16)–(19) są takie jak w poprzednich modelach.

Przedstawiony powyżej model posłużył do rozwiązania przykładu. Uzyskano następujące wyniki. Wartość funkcji celu wyniosła $z = 38,1$, co oznacza sumę zdyskontowanych przepływów pieniężnych poszczególnych czynności na zakończeniu projektu. Projekt kończy się po 10 jednostkach czasu, a więc zgodnie z czasem wyznaczonym za pomocą metody ścieżki krytycznej, co widoczne jest na harmonogramie projektu (rys. 9), przy czym czynności umieszczone zostały na harmonogramie w najwcześniejszych czasach rozpoczęcia i zakończenia (zgodnie z czasami z CPM). Wykorzystanie zasobów nie przekroczyło założonego w przykładzie poziomu 5 jednostek (rys. 10).



Rys. 9. Maksymalizacja zdyskontowanych przepływów pieniężnych – harmonogram
 Źródło: opracowanie własne.



Rys. 10. Maksymalizacja zdyskontowanych przepływów pieniężnych – poziom wykorzystania zasobów
 Źródło: opracowanie własne.

3.5. Wielokryterialne harmonogramowanie projektów jako zadanie programowania zero-jedynkowego

Jak dowiedziono powyżej, możliwe jest budowanie modeli matematycznych harmonogramowania projektów, w których optymalizacji podlegają takie czynniki, jak: czas trwania projektu, poziom wykorzystania zasobów czy zdyskontowane przepływy pieniężne. Ponieważ w praktyce często zdarza się, że istotne jest ustalenie harmonogramu optymalnego nie tylko ze względu na jeden czynnik, warto rozważyć połączenie wskazanych wcześniej modeli w celu zbudowania wielokryterialnego modelu harmonogramowania projektów. Taki model będzie miał wówczas następującą postać:

$$\sum_{t=1}^T \sum_{j=1}^J \max \{ \{0, t - LF_j\} \} \cdot x_{jt} \rightarrow \min \quad (j = 1, \dots, J, t = 1, \dots, T), \quad (20)$$

$$\max_{t=1, \dots, T} \left[\sum_{j=1}^J r_{jk}^r \cdot x_{jt} \right] \rightarrow \min \quad (j = 1, \dots, J, t = 1, \dots, T, k = 1, \dots, K), \quad (21)$$

$$\sum_{i=1}^J \left[\sum_{t=1}^{d_j} [c_{jt}^r \cdot e^{\alpha(d_j-t)}] \cdot e^{-\alpha F_j} \right] \rightarrow \max \quad (j = 1, \dots, J, t = 1, \dots, T), \quad (22)$$

przy ograniczeniach:

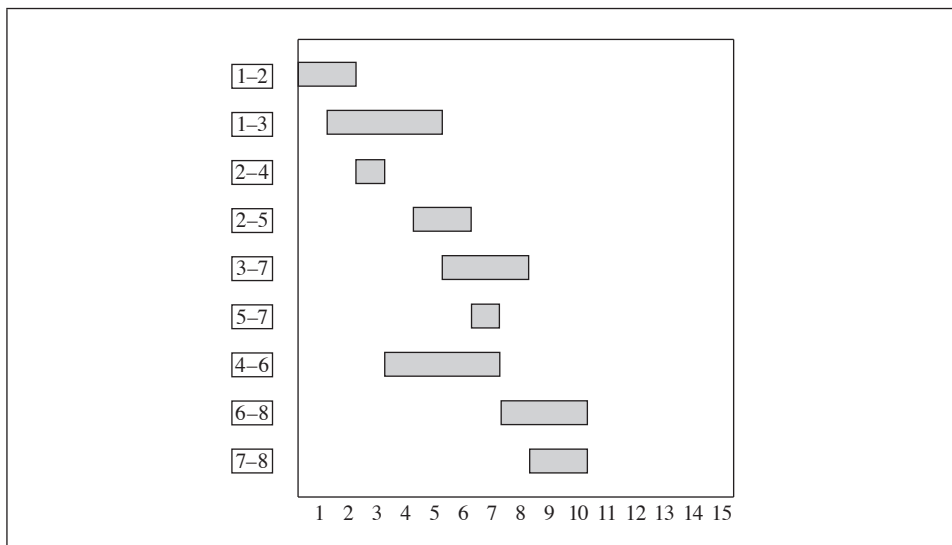
$$F_j - d_j \geq F_{ij} \quad (j = 1, \dots, J, t = 1, \dots, T, i \in S_{ij}), \quad (23)$$

$$\sum_{j=1}^J r_{jk} \cdot x_{jt} \leq R_{kt} \quad (j = 1, \dots, J, t = 1, \dots, T, k = 1, \dots, K), \quad (24)$$

$$\sum_{t=1}^T x_{jt} = 1 \quad (j = 1, \dots, J, t = 1, \dots, T), \quad (25)$$

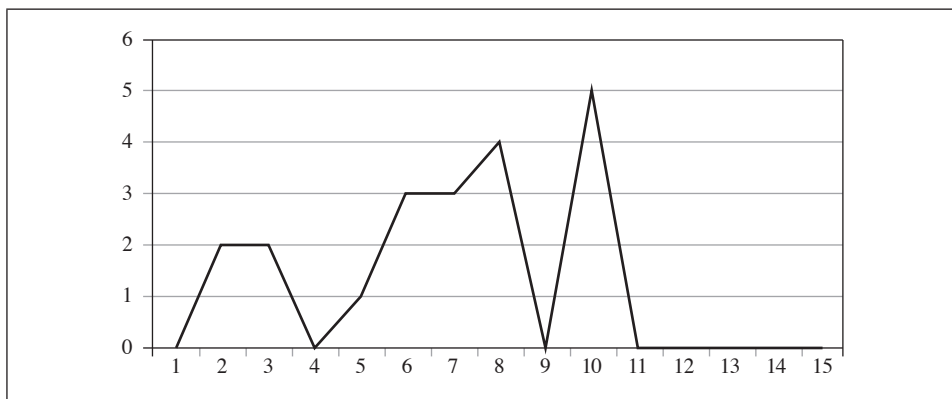
$$x_{jt} \in \{0, 1\} \quad (j = 1, \dots, J, t = 1, \dots, T). \quad (26)$$

Powyższy model wielokryterialny posłużył do rozwiązania przykładu. Dla uproszczenia jako metodę rozwiązania problemu wykorzystano metodę agregacji funkcji celu. Wynika to z faktu, że zastosowane funkcje celu są nieliniowe, co uniemożliwia zastosowanie klasycznych metod rozwiązywania problemów wielokryterialnych. Przyjęto następujące wagi dla poszczególnych funkcji kryterium: 40% dla optymalizacji czasu trwania projektu, 40% dla optymalizacji poziomu wykorzystania zasobów oraz 20% dla optymalizacji przepływów pieniężnych. Uzyskano harmonogram, w którym projekt kończy się po 10 jednostkach czasu (rys. 11), a poziom wykorzystania zasobów nie przekroczył zakładanego poziomu 5 jednostek (rys. 12).



Rys. 11. Wielokryterialne harmonogramowanie projektu – harmonogram

Źródło: opracowanie własne.



Rys. 12. Wielokryterialne harmonogramowanie projektów – poziom wykorzystania zasobów

Źródło: opracowanie własne.

4. Podsumowanie

Zastosowanie metod optymalizacyjnych umożliwia stworzenie harmonogramów projektów optymalnych nie tylko ze względu na czas, ale również zasoby czy wskaźniki finansowe.

W niniejszej pracy przedstawione zostały modele optymalizacji harmonogramów projektów wykorzystujące zmienną zero-jedynkową. Głównymi zaletami przedstawionego podejścia są jego prostota oraz łatwość zastosowania w wielu metodach optymalizacyjnych ze względu na zero-jedynkową postać zmiennej. Wadą przedstawionego modelu jest liczba zmiennych, która wynosi $j \times p \times t$. Wraz ze wzrostem liczby czynności czy horyzontu planowania projektu liczba zmiennych będzie wzrastać wykładniczo, co może utrudniać rozwiązanie większych problemów.

Dzięki binarnej postaci zmiennej decyzyjnej możliwe jest zastosowanie algorytmów genetycznych do rozwiązania wielokryterialnego problemu harmonogramowania projektów. Ponieważ problem ten jest problemem NP-trudnym, zastosowanie tych algorytmów jest zasadne. Będzie ono celem dalszych badań w tym zakresie.

W dalszych badaniach rozważona powinna zostać również możliwość zastosowania sieci deterministycznej stopnia pierwszego *Activity on Node*, będącej acyklicznym grafem skierowanym $G = (W, K)$, w którym W to zbiór wierzchołków reprezentujących czynności, a K to zbiór łuków prezentujących relacje kolejnościowe. Sieć ta, zwana także siecią czynności, umożliwia zastosowanie nie tylko relacji kolejnościowych typu zakończenie–początek, ale także relacji typu: zakończenie–zakończenie, rozpoczęcie–rozpoczęcie czy rozpoczęcie–zakończenie.

Literatura

- Bartusch M., Mohring R., Radermacher F. [1988], *Scheduling Project Networks with Resource Constraints and Time Windows*, „Annals of Operations Research”, vol. 16, nr 1–4.
- Bianco L., Dell’Olmo P., Speranza M. [1998], *Heuristics for Multimode Scheduling Problem with Dedicated Resources*, „European Journal of Operational Research”, vol. 107.
- Brandenburg H. [2002], *Zarządzanie projektami*, Akademia Ekonomiczna w Katowicach, Katowice.
- Doersch R., Patterson J. [1977], *Scheduling a Project to Maximize Its Present Value: A Zero-one Programming Approach*, „Management Science”, vol. 23, nr 8.
- Icmeli O., Erenguc S. [1996], *A Branch and Bound Procedure for the Resource Constrained Project Scheduling Problem with Discounted Cash Flows*, „Management Science”, vol. 42, nr 10.

- Kostrubiec A. [2003], *Harmonogramowanie projektów – przegląd modeli* [w:] *Inżynieria zarządzania przedsięwzięciami*, red. L. Zawadzka, Wydawnictwo Politechniki Gdańskiej, Gdańsk.
- Norberciak M. [2002], *Przegląd metod automatycznego planowania – przykład wykorzystania algorytmu genetycznego w rozwiązaniu prostego problemu planowania*, Prace Naukowe Wydziałowego Zakładu Informatyki Politechniki Wrocławskiej, z. 1: Sztuczna inteligencja, Wrocław.
- Russell R. [1970], *Cash Flows in Networks*, „Management Science”, vol. 16, nr 5.
- Shouman M. i in. [2006], *Genetic Algorithm Constraint Project Scheduling*, „Alexandria Engineering Journal”, vol. 45, nr 3.
- Talbot T. [1982], *Resource-constrained Project Scheduling with Time-resource Tradeoffs: The Non-preemptive*, „Management Science”, vol. 28, nr 10.
- Vanhoucke M., Demeulemeester E., Herroelen W. [2001], *On Maximizing the Net Present Value of a Project under Renewable Resource Constraints*, „Management Science”, vol. 47, nr 8.
- Viana A., de Sousa J. [2000], *Using Metaheuristic in Multiobjective Resource Constrained Project Scheduling*, „European Journal of Operational Research”, vol. 120, nr 20.

Using the Zero-one Programming Approach in Project Scheduling

Because of increasing interest in project management in the subject literature, optimisation techniques are often considered in project planning and scheduling. The main project scheduling techniques are CPM or PERT. Those methods deliver schedules with optimal project finish times and ensure the level of resource usage is controlled. In practical situations, project schedules should be optimised not only because of time but also because of resource usage and cash flows. The above techniques do not meet those requirements.

The purpose of this paper is to demonstrate how zero-one programming can be used in project scheduling. Zero-one programming is a special case of integer programming where all the decision variables are integers and can assume values of either zero or one. A zero-one programming formulation has been applied to solve three project scheduling problems, namely the optimisation of project completion time, resource usage and project cash flows.

Keywords: project scheduling, zero-one programming, schedule optimisation, multiple criteria optimisation.

Lista recenzentów Zeszytów Naukowych Uniwersytetu Ekonomicznego w Krakowie opublikowanych w 2013 roku

Jacek Batóg
Jan L. Bednarczyk
Henryk Bieniok
Andrzej Buszko
Bożena Borycka
Szymon Cyfert
Janina Czapska
Bolesław Domański
Ewa Drabik
Marek Drwięga
Dorota Dąbek
Joanna Ejdys
Teresa Famulska
Beata Zofia Filipiak
Piotr Fiszeder
Stanisław Flejterski
Eugeniusz Gatnar
Józef Garczarczyk
Krzysztof Gawroński
Robert Geisler
Zygmunt Górka
Wacław Jarmołowicz
Teresa Kamińska
Jacek Klich
Danuta Kołożyn-Krajewska
Witold Kazimierz Krajewski
Anna Kostur
Marianna Księżyk
Miroslaw Krzyśko

Hieronim Kubera
Elżbieta Kucka
Joanna Lach
Ján Lisý
Elżbieta Lorek
Grzegorz Maciejewski
Alicja Maleszka
Małgorzata Markowska
Teresa Mróz
Wiesław Musiał
Joanicjusz Nazarko
Janina Pach
Krystyna Rajzer
Wojciech Rybicki
Małgorzata Rószkiewicz
Grażyna Skąpska
Elżbieta Sobczak
Łukasz Tomasz Sroka
Danuta Stawasz
Janusz Sztumski
Ryszard Szewczyk
Maria Śmiechowska
Andrzej Świątkowski
Anna Walaszek-Pyziół
Tadeusz Wawak
Dorota Witkowska
Tadeusz Włudyka
Michał Woźniak
Feliks Wysocki
Urszula Zagóra-Jonszta
Jan Zawadzki
Zofia Żakowska
Janusz Żmija
Jerzy Żuchowski