POLSKA AKADEMIA NAUK
KOMITET STATYSTYKI I EKONOMETRII

PRZEGLĄD STATYSTYCZNY

STATISTICAL REVIEW

TOM 58

1-2

2011

WARSZAWA 2011

WYDAWCA

Komitet Statystyki i Ekonometrii Polskiej Akademii Nauk

Wydanie publikacji dofinansowane przez Ministra Nauki i Szkolnictwa Wyższego

Strona WWW „Przegląd Statystyczny”:
http://www.przegladstatystyczny.pan.pl

Nakład: 300 egz.

PAN Warszawska Drukarnia Naukowa
00-056 Warszawa, ul. Śniadeckich 8
Tel./fax:+48 22 628 76 14

Od Redakcji

Pragniemy wyrazić ubolewanie z powodu opóźnienia w ukazaniu się pierwszych tegorocznych numerów „Przeglądu Statystycznego". Niezawinionym przez nas powodem tego opóźnienia była przedłużająca się procedura, mająca na celu wyłonienie podmiotu odpowiedzialnego za druk kwartalnika. Ostatecznie obowiązki te od bieżącego roku po firmie ELIPSA przejęła Polska Akademia Nauk Warszawska Drukarnia Naukowa. Mamy nadzieję, że współpraca z PAN WDN gwarantować będzie nadal wysoki poziom edytorski tekstów zamieszczanych na łamach „Przeglądu Statystycznego".

Kolejny numer 3-4 z roku 2011 ukaże się pod koniec bieżącego roku. Zachęcamy gorąco wszystkich zainteresowanych do nadsyłania tekstów do „Przeglądu Statystycznego" i obiecujemy, że kolejny rok kwartalnik nasz rozpocznie bez opóźnienia.

Mirosław Szreder – Redaktor Naczelny
Magdalena Osińska – Zastępca Redaktora Naczelnego
Marek Walesiak – Zastępca Redaktora Naczelnego
Krzysztof Najman – Sekretarz Naukowy Redakcji

PRZEGLĄD STATYSTYCZNY
R. LVIII – ZESZYT 1-2 – 2011

ALEKSANDRA ŁUCZAK, FELIKS WYSOCKI

PORZĄDKOWANIE LINIOWE OBIEKTÓW Z WYKORZYSTANIEM ROZMYTYCH METOD AHP I TOPSIS[1]

1. WPROWADZENIE

Porządkowanie liniowe jest procesem badawczym, który prowadzi do ustalenia rankingu obiektów z punktu widzenia przyjętego kryterium porządkowania. W procesie tym wykorzystuje się wartości cechy syntetycznej – syntetycznego miernika rozwoju. Stanowią one oceny obiektów, według których następuje ich porządkowanie w kolejności od najlepszego do najgorszego.

Koncepcję budowy cechy syntetycznej stworzył Hellwig [5] przed ponad czterdziestu laty. Była ona „źródłem inspiracji dla kolejnych pokoleń statystyków i ekonometryków, podejmujących liczne próby modyfikacji i wysuwających kolejne propozycje metod porządkowania liniowego"[17] (s. 145).

W każdym etapie procedury budowy cechy syntetycznej pojawiają się pytania dotyczące sposobu postępowania. W pracy zaproponowano rozwiązania metodyczne, które poszerzają możliwości porządkowania liniowego obiektów i rozpoznawania typów rozwojowych. Dotyczą one uwzględniania w procesie agregacji zarówno cech metrycznych, jak i niemetrycznych (mierzonych na skali porządkowej) a także budowy systemu wag cech.

Współczynniki wagowe mogą być ustalone przy zastosowaniu dwóch grup metod. Pierwsza wykorzystuje procedury statystyczne, druga zaś opiera się na opiniach ekspertów [por. 6]. Podejście statystyczne bazuje na informacjach o cechach tkwiących tylko w samej macierzy danych, a w szczególności wykorzystuje analizę zmienności cech oraz analizę korelacji między cechami albo tylko jedną z tych analiz. Jego specyfiką jest „mechaniczne potraktowanie problemu ważenia, abstrahujące od rzeczywistej pozycji danej cechy określonej przesłankami merytorycznymi" [6] (s. 70). Oznacza to, że cechy o dużej zmienności nie muszą być wcale ważne w sensie merytorycznym. Dlatego przy ustalaniu wag cech właściwszym podejściem wydaje się zastosowanie metody ekspertów.

---

[1] Artykuł jest rozszerzoną wersją publikacji: *Wykorzystanie rozmytych metod AHP i TOPSIS do porządkowania liniowego obiektów* zamieszczonej w Pracach Naukowych Uniwersytetu Ekonomicznego we Wrocławiu – *Taksonomia 17. Klasyfikacja i analiza danych – teoria i zastosowania* [7].

Celem niniejszej pracy jest przedstawienie możliwości wykorzystania w procesie tworzenia cechy syntetycznej rozmytego analitycznego procesu hierarchicznego (*Fuzzy Analytic Hierarchy Process* – FAHP) i rozmytej metody TOPSIS (*Technique for Order Preference by Similarity to an Ideal Solution*). FAHP wykorzystuje opinie ekspertów do ustalania współczynników wagowych określających ważność cech oraz jednocześnie pozwala na eliminację cech o najmniejszym znaczeniu w zagadnieniu porządkowania liniowego obiektów [2]. W tym przypadku wagi cech ustala się na podstawie rozmytych opinii ekspertów, tzw. miękkich opinii (*soft opinions*), które są bardziej realistyczne aniżeli opinie dokładne (*hard opinions*). Wykorzystuje się je następnie w procesie tworzenia cechy syntetycznej za pomocą rozmytej metody TOPSIS [4, 17], która jest procedurą statystyczną prowadzącą do porządkowania liniowego obiektów opisanych za pomocą cech metrycznych i niemetrycznych – porządkowych[2]. Zatem w artykule przedstawiono zintegrowane podejście do wyznaczania cechy syntetycznej obejmujące zastosowanie rozmytych metod FAHP i TOPSIS.

Procedurę tworzenia cechy syntetycznej zastosowano do oceny poziomu rozwoju społeczno-gospodarczego powiatów województwa wielkopolskiego.

2. METODYKA BADAŃ

W procesie tworzenia cechy syntetycznej bazującym na rozmytych metodach AHP i TOPSIS można wyróżnić następujące etapy postępowania:

Etap 1. Utworzenie struktury hierarchicznej wielokryterialnego problemu oceny obiektów.

Etap 2. Określenie ważności kryteriów i cech poprzez przyporządkowanie im współczynników wagowych uzyskanych z rozmytego analitycznego procesu hierarchicznego (FAHP).

Etap 3. Wyznaczenie wartości cechy syntetycznej (syntetycznego miernika rozwoju) za pomocą rozmytej metody TOPSIS.

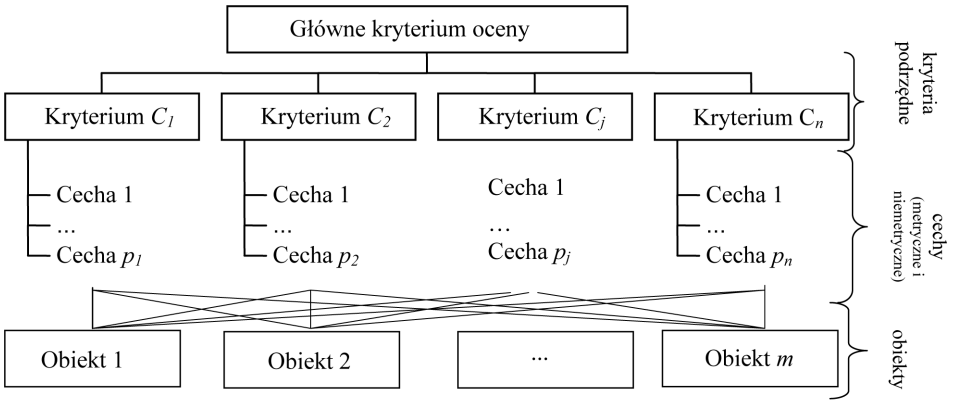
Etap 4. Uporządkowanie liniowe i klasyfikacja typologiczna obiektów według wartości cechy syntetycznej.

Etap 1. Struktura hierarchiczna wielokryterialnego problemu oceny obiektów jest tworzona drogą rozkładu rozważanego problemu na elementy składowe: główne kryterium oceny (np. poziom rozwoju społeczno-gospodarczego), kryteria podrzędne, cechy proste oraz oceniane obiekty (rys. 1).

Kryterium główne i kryteria podrzędne oraz cechy opisujące badane obiekty są wzajemnie powiązane. Wybór kryteriów i cech powinien opierać się na przesłankach merytorycznych i statystycznych. Dane statystyczne można zapisać w postaci macierzy $\mathbf{X}_{Cj} = \{x_{ik}, i= 1, 2, \ldots, m; k= 1, 2, \ldots, p_j\}$, gdzie m jest liczbą obiektów, p_j jest liczbą cech w ramach kryterium w_j ($j = 1, 2, \ldots, n$), $p_1 + p_2 + ... + p_n = P$ jest łączną liczbą cech.

[2] W procedurze budowy cechy syntetycznej można również uwzględniać cechy metryczne i porządkowe stosując miarę odległości GDM2 zaproponowaną przez M. Walesiaka [12]

Etap 2. Określenie systemu wag dla kryteriów $\mathbf{W} = (w_1, w_2, ..., w_n)$ i cech $\mathbf{W}_j = \left(w_{j1}, w_{j2}, ..., w_{jp_j}\right)$ $(j = 1, 2, ..., n;\ k = 1, 2, ..., p_j)$. Wektory współczynników wagowych można otrzymać metodą rozmytego analitycznego procesu hierarchicznego zaproponowaną przez Changa [2]. Metoda składa się z następujących kroków [zob. 2, 14]:



Rysunek 1. Struktura hierarchiczna wielokryterialnego problemu oceny obiektów

Źródło: Opracowanie własne.

Krok 1. Porównanie parami cech w ramach kryterium oceny. Dokonuje się porównań parami ważności cech w odniesieniu do danego kryterium podrzędnego wykorzystując do tego np. dziewięciostopniową skalę Saaty'ego (tab. 1). Wyniki porównań są przedstawiane w postaci rozmytych macierzy porównań parami $\tilde{\mathbf{A}}_j$:

$$\tilde{\mathbf{A}}_j = \left[\tilde{a}_{jkg}\right] = \begin{bmatrix} (1,1,1) & \left(l_{j12}, m_{j12}, u_{j12}\right) & ... & \left(l_{j1p_j}, m_{j1p_j}, u_{j1p_j}\right) \\ \left(l_{j21}, m_{j21}, u_{j21}\right) & (1,1,1) & ... & \left(l_{j2p_j}, m_{j2p_j}, u_{j2p_j}\right) \\ ... & ... & ... & ... \\ \left(l_{jp_j1}, m_{jp_j1}, u_{jp_j1}\right) & l_{jp_j2}, m_{jp_j2}, u_{jp_j2} & ... & (1,1,1) \end{bmatrix}, \qquad (1)$$

gdzie: $\tilde{a}_{jkg} = \left(l_{jkg}, m_{jkg}, u_{jkg}\right)$ i $\tilde{a}_{jgk} = \tilde{a}_{jkg}^{-1} = \left(1/u_{jkg}, 1/m_{jkg}, 1/l_{jkg}\right)$, $(j= 1, 2, \ldots, n)$; $k, g = 1, 2, \ldots p_j$, oraz $k \neq g$, $\tilde{a}_{jkg}$ są ocenami porównań parami cech w ramach j-tego kryterium określonymi przez ekspertów lub średnimi z ocen grupy ekspertów.

Krok 2. Wyznaczenie sumy wartości elementów każdego wiersza rozmytej macierzy porównań parami $\tilde{\mathbf{A}}_j$ $(j= 1, 2, \ldots, n)$ *i normalizacja sum wierszowych za pomocą operacji na liczbach rozmytych:*

$$\tilde{Q}_{jk} = \left(l_{jk}, m_{jk}, u_{jk}\right) = \sum_{g=1}^{p_j} \left(l_{jkg}, m_{jkg}, u_{jkg}\right) \otimes \left[\sum_{k=1}^{p_j} \sum_{g=1}^{p_j} \left(l_{jkg}, m_{jkg}, u_{jkg}\right)\right]^{-1}, \qquad (2)$$

$$\mathrm{j} = 1, 2, \ldots, n; \quad \mathrm{k} = 1, 2, \ldots, p_j.$$

Tabela 1

Dziewięciostopniowa skala preferencji między dwoma porównywanymi elementami według Saaty'ego

Przewaga ważności cech (kryteriów)	Preferencje opisane słownie	Siła przewagi ważności	
		klasyczna AHP	rozmyta AHP
Równoważność	Oba elementy (cechy, kryteria) przyczyniają się równo do osiągnięcia celu (jeden element ma takie samo znaczenie jak drugi)	1	$\tilde{1} = (1, 1, 1)$
Słaba lub umiarkowana	Nie przekonywujące znaczenie lub słaba preferencja jednego elementu nad drugim (jeden element ma nieco większe znaczenie niż drugi)	3	$\tilde{3} = (1, 3, 5)$
Istotna, zasadnicza, mocna	Zasadnicze lub mocne znaczenie lub mocna preferencja jednego elementu nad innymi (jeden element ma wyraźnie większe znaczenie niż drugi)	5	$\tilde{5} = (3, 5, 7)$
Zdecydowana lub bardzo mocna	Zdecydowane znaczenie lub bardzo mocna preferencja jednego elementu nad innym (jeden element ma bezwzględnie większe znaczenie niż drugi)	7	$\tilde{7} = (5, 7, 9)$
Absolutna	Absolutne znaczenie lub absolutna preferencja jednego elementu nad innym	9	$\tilde{9} = (7, 9, 9)$
Dla porównań kompromisowych pomiędzy powyższymi wartościami	Czasami istnieje potrzeba interpolacji numerycznej kompromisowych opinii, ponieważ nie ma odpowiedniego słownictwa do ich opisania, przeto stosujemy pośrednie wartości między dwoma sąsiednimi ocenami	2, 4, 6 i 8	$\tilde{2} = (1, 2,4)$; $\tilde{4} = (2, 4, 6)$; $\tilde{6} = (4, 6, 8)$; $\tilde{8} = (6, 8, 9)$
Przechodniość ocen	Jeżeli i-ty element ma przypisany jeden z powyższych stopni podczas porównania do j-tego elementu, wtedy j-ty element ma odwrotną wartość, gdy porównuje się do i-tego (jeżeli porównując X z Y przyporządkowujemy wartość α, to wtedy automatycznie musimy przyjąć, że wynikiem porównania Y z X musi być $1/\alpha$)	odwrotności powyższych wartości	odwrotności powyższych wartości

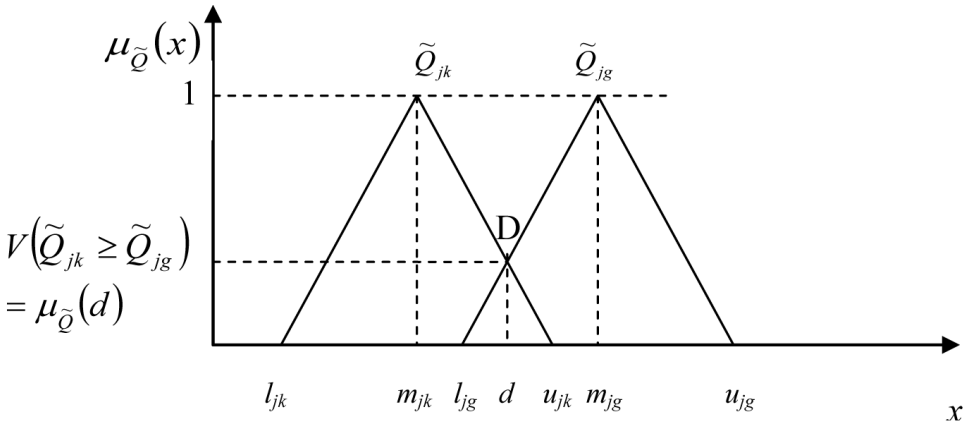
Źródło: Opracowanie własne na podstawie [10, 13].

Krok 3. Obliczenie stopnia możliwości, że liczba rozmyta $\tilde{Q}_{jk}$ jest większa bądź równa liczbie $\tilde{Q}_{jg}$, czyli że $\tilde{Q}_{jk} \geqslant \tilde{Q}_{jg}$ za pomocą następującego równania:

$$V\left(\tilde{Q}_{jk} \geqslant \tilde{Q}_{jg}\right) = \begin{cases} 1, \text{ dla } m_{jk} \geqslant m_{jg} \\ 0, \text{ dla } l_{jg} \geqslant u_{jk} \\ \dfrac{l_{jg} - u_{jk}}{(m_{jk} - u_{jk}) - (m_{jg} - l_{jg})} \text{ w innych przypadkach,} \end{cases} \tag{3}$$

gdzie $\tilde{Q}_{jk} = \left(l_{jk}, m_{jk}, u_{jk},\right)$ i $\tilde{Q}_{jg} = \left(l_{jg}, m_{jg}, u_{jg}\right)$ są dwiema liczbami rozmytymi.

Rys. 2 ilustruje równanie $V\left(\tilde{Q}_{jk} \geqslant \tilde{Q}_{jg}\right) = \mu_{\tilde{Q}}(d)$ dla przypadku $m_{jk} < l_{jg} < u_{jk} < m_{jg}$, gdzie d jest odciętą korespondującą z punktem przecięcia $D = (d, \mu(d))$ dwóch trójkątnych funkcji przynależności $\mu_{\tilde{Q}_{jk}}(x)$ i $\mu_{\tilde{Q}_{jg}}(x)$. Porównując $\tilde{Q}_{jk}$ i $\tilde{Q}_{jg}$ trzeba wyznaczyć zarówno $V\left(\tilde{Q}_{jk} \geqslant \tilde{Q}_{jg}\right)$, jak i $V\left(\tilde{Q}_{jg} \geqslant \tilde{Q}_{jk}\right)$.



Rysunek 2. Wyznaczenie współrzędnych punktu przecięcia między $\tilde{Q}_{jk}$ i $\tilde{Q}_{jg}$

Źródło: Opracowanie własne na podstawie Chang [2].

Krok 4. Wyznaczenie najmniejszego stopnia możliwości $V\left(\tilde{Q}_{jk} \geqslant \tilde{Q}_{jg}\right)$ liczby rozmytej $\tilde{Q}_{jk}$ względem wszystkich pozostałych $(p_j - 1)$ liczb rozmytych jako:

$$V\left(\tilde{Q}_{jk} \geq \tilde{Q}_{jg} \middle| g = 1, ..., p_j\,; k \neq g\right) = \min_{\substack{g \in (1,...,p_j) \\ g \neq k}} V\left(\tilde{Q}_{jk} \geq \tilde{Q}_{jg}\right) = \\ = \mu_{\tilde{Q}_{jk}}(d) = \mu_{\tilde{Q}_{jg}}(d)\,; \mathrm{k} = 1, 2, \ldots, p_j. \tag{4}$$

Krok 5. Obliczenie wskaźników udziału:

$$w_{jk}^{(l)} = \frac{V\left(\tilde{Q}_{jk} \geqslant \tilde{Q}_{jg} \middle| g = 1, ..., p_j\,; k \neq g\right)}{\sum_{h=1}^{p_j} V\left(\tilde{Q}_{jh} \geqslant \tilde{Q}_{jg} \middle| g = 1, ..., p_j\,; h \neq g\right)}; \quad k = 1, 2, \ldots, p_j, \tag{5}$$

które przyjmowane są jako wagi lokalne[3] cechy.

Krok 6. Obliczenie wartości globalnych współczynników wagowych[4]. Oblicza się je mnożąc lokalne współczynniki wagowe przez współczynniki wagowe dla kryteriów

[3] Wagi lokalne określają względną ważność cech w ramach danego kryterium podrzędnego. Suma wag lokalnych dla cech w ramach każdego kryterium podrzędnego wynosi 1.

[4] Wagi globalne cech reprezentują ważność cech w odniesieniu do kryterium głównego. Suma wszystkich wag globalnych dla cech wynosi 1.

$w_{jk} = w_{jk}^{(l)} \cdot w_j$. W rezultacie wielkości w_{jk} przyjmuje się jako współczynniki wagowe dla cech i przedstawia w postaci wektora $\mathbf{W}_j = \left(w_{j1}, w_{j2}, ..., w_{jp_j}\right)^{\mathrm{T}}$, przy czym $\sum_{k=1}^{p_j} w_{jk} = w_j$, $\sum_{j=1}^{n} w_j = 1$, $\underset{j}{\forall} w_j \geqslant 0$.

Analogiczne według kroków 1-5 można obliczyć wagi w_j dla kryteriów, przy czym lokalne i globalne współczynniki wagowe dla danego kryterium są identyczne.

Etap 3. Wyznaczenie wartości cechy syntetycznej S_i za pomocą rozmytej metody TOPSIS. Najpierw wartości cech zamienią się na trójkątne liczby rozmyte $\tilde{x} = (a, b, c)$. W zależności od typu danych stosuje się przekształcenia:

a) cechy metryczne

– dane punktowe $\left(x_{ik}, i = 1, 2, \ldots, m,\ k = 1, 2, \ldots, p_j\right)$ sprowadza się do liczb rozmytych przyjmując $\tilde{x}_{ik} = (x_{ik}, x_{ik}, x_{ik}) = (b, b, b)$,

– dane przedziałowe $\left(x_{ik} \in \left[(x_{ik})^L ; (x_{ik})^U\right]\right)$ przekształca się przyjmując $\tilde{x}_{ik} = (a, b, c)$, gdzie $a = (x_{ik})^L$, b = wartość średnia, $c = (x_{ik})^U$,

b) cechy porządkowe – wartości cech sprowadza się do poziomów zmiennej lingwistycznej, którym odpowiadają trójkątne liczby rozmyte $\tilde{x}_{ik} = (a, b, c)$, reprezentowane przez trzy oceny: pesymistyczną, najbardziej prawdopodobną i optymistyczną (tab. 2).

Tabela 2

Zmienna lingwistyczna – jej poziomy i odpowiadające im liczby rozmyte

Poziomy zmiennej lingwistycznej	bardzo niski (BN)	niski (N)	średni (Ś)	wysoki (W)	bardzo wysoki (BW)
Trójkątne liczby rozmyte (a, b, c)	(0, 0, 20)	(20, 30, 40)	(40, 50, 60)	(60, 70, 80)	(80, 100, 100)

Źródło: [3].

Wszystkie relacje między liczbami rozmytymi można przedstawić za pomocą działań na ich ocenach (zob. *Dodatek*). Otrzymane trójkątne liczby rozmyte są przedstawiane w postaci rozmytych macierzy danych $\tilde{\mathbf{X}}_{\mathrm{C_j}}$=$\{\tilde{x}_{ik} = (a_{ik}, b_{ik}, c_{ik}),\ i = 1, 2, ..., m;\ k = 1, ..., p_j\}$.

Kolejnym krokiem procedury jest normalizacja liczb rozmytych, mająca na celu ujednolicenie rzędów ich wielkości. Jednocześnie destymulany oraz nominanty zostają przekształcone na stymulanty. Stosuje się przekształcenie ilorazowe wykonywane na liczbach rozmytych [4]:

– dla stymulant

$$\tilde{z}_{ik} = \left(\frac{a_{ik}}{c_k^+}, \frac{b_{ik}}{c_k^+}, \frac{c_{ik}}{c_k^+}\right), \tag{6}$$

gdzie: $c_k^+ = \max_i c_{ik}$, $\max_i c_{ik} \neq 0$,

– dla destymulant

$$\tilde{z}_{ik} = \left(\frac{a_k^-}{c_{ik}}, \frac{a_k^-}{b_{ik}}, \frac{a_k^-}{a_{ik}}\right), \text{ gdy } a_{ik}^{'} b_{ik}^{'} c_{ik}^{\neq} 0, \tag{7}$$

gdzie: $a_k^- = \min\limits_i a_{ik}$,

– dla nominant

$$\tilde{z}_{ik} = \left(\frac{a_{ik}}{b_k^*}, \frac{b_{ik}}{b_k^*}, \frac{c_{ik}}{b_k^*}\right), \text{ gdy } b_{ik} \leqslant \text{nom } b_{ik}, \text{ gdzie} : b_k^* = \operatorname{nom}_i b_{ik}, \text{nom } b_{ik} \neq 0, \tag{8}$$

$\tilde{z}_{ik} = \left(\frac{b_k^*}{c_{ik}}, \frac{b_k^*}{b_{ik}}, \frac{b_k^*}{a_{ik}}\right)$, gdy $b_{ik} > \text{nom } b_{ik}$, $a_{ik}^{'} b_{ik}^{'} c_{ik}^{\neq} 0$, $i = 1, 2, ..., m$; $k = 1, ..., p_j$.

Znormalizowane wartości cech z każdej macierzy $\tilde{\mathbf{Z}}_{\text{Cj}}$=$\{\tilde{z}_{ik} = (\alpha_{ik}, \beta_{ik}, \gamma_{ik})$, $i = 1, 2, ..., m$; $k = 1, ..., p_j\}$ zestawia się w jedną macierz $\tilde{\mathbf{Z}}$ =$\{\tilde{z}_{ik} = (\alpha_{ik}, \beta_{ik}, \gamma_{ik})$, $i = 1, 2, ..., m$; $k = 1, ..., P\}$, gdzie $p_1 + p_2 + ... + p_n = P$.

Znormalizowane liczby rozmyte pomnożone przez współczynniki wagowe ważności cech można zapisać w postaci macierzy:

$$\tilde{\mathbf{R}} = [\tilde{r}_{ik}]_{m \times P}, \text{ przy czym } \tilde{r}_{ik} = \tilde{z}_{ik} \otimes w_k \quad i = 1, 2, \ldots, m; k = 1, 2, \ldots, P. \tag{9}$$

Następnie ustalone zostają współrzędne obiektów modelowych – rozmytego wzorca $\tilde{A}^+$ i antywzorca rozwoju $\tilde{A}^-$:

$$\tilde{A}^+ = \left(\max_i (\tilde{r}_{i1}), \max_i (\tilde{r}_{i2}), ..., \max_i (\tilde{r}_{iP})\right) = (\tilde{r}_1^+, \tilde{r}_2^+, ..., \tilde{r}_P^+) \tag{10}$$

$$\tilde{A}^- = \left(\min_i (\tilde{r}_{i1}), \min_i (\tilde{r}_{i2}), ..., \min_i (\tilde{r}_{iP})\right) = (\tilde{r}_1^-, \tilde{r}_2^-, ..., \tilde{r}_P^-) \tag{11}$$

i na tej podstawie obliczone zostają odległości każdego ocenianego obiektu od wzorca A^+ i antywzorca rozwoju A^-:

$$d_i^+ = \sum_{k=1}^{P} d\left(\tilde{r}_{ik}; \tilde{r}_k^+\right), \quad d_i^- = \sum_{k=1}^{P} d\left(\tilde{r}_{ik}; \tilde{r}_k^-\right), (i = 1, 2, \ldots, m), \tag{12}$$

Odległość między dwiema trójkątnymi liczbami rozmytymi $\tilde{x}_1 = (a_1, b_1, c_1)$ i $\tilde{x}_2 = (a_2, b_2, c_2)$ jest zdefiniowana następująco (Chen [4]):

$$d(\tilde{x}_1; \tilde{x}_2) = \sqrt{\frac{1}{3}\left((a_1 - a_2)^2 + (b_1 - b_2)^2 + (c_1 - c_2)^2\right)}. \tag{13}$$

W kolejnym kroku oblicza się wartości syntetycznego miernika rozwoju:

$$S_i = \frac{d_i^-}{d_i^+ + d_i^-}, \quad 0 \leqslant S_i \leqslant 1, (i = 1, 2, \ldots, m). \tag{14}$$

Im mniejsza jest odległość danego obiektu od obiektu modelowego – wzorca rozwoju, a tym samym większa od drugiego bieguna – antywzorca rozwoju, tym wartość miernika syntetycznego jest bliższa 1.

Etap 4. Uporządkowanie liniowe i klasyfikacja typologiczna obiektów według wartości cechy syntetycznej zgodnie z następującymi zasadami [8, 9]:

klasa I (poziom bardzo wysoki): $S_i \geqslant \bar{S} + 2 \cdot S_c$
klasa II (wysoki): $\bar{S} + S_c \leqslant S_i < \bar{S} + 2 \cdot S_c$
klasa III (średni-wyższy) $\bar{S} \leqslant S_i < \bar{S} + S_c$ (15)
klasa IV (średni-niższy): $\bar{S} - S_c \leqslant S_i < \bar{S}$
klasa V (niski): $\bar{S} - 2 \cdot S_c \leqslant S_i < \bar{S} - S_c$
klasa VI (bardzo niski): $S_i < \bar{S} - 2 \cdot S_c$

gdzie: $\bar{S}$ – jest średnią arytmetyczną z wartości cechy syntetycznej, a S_c – ich odchyleniem standardowym.

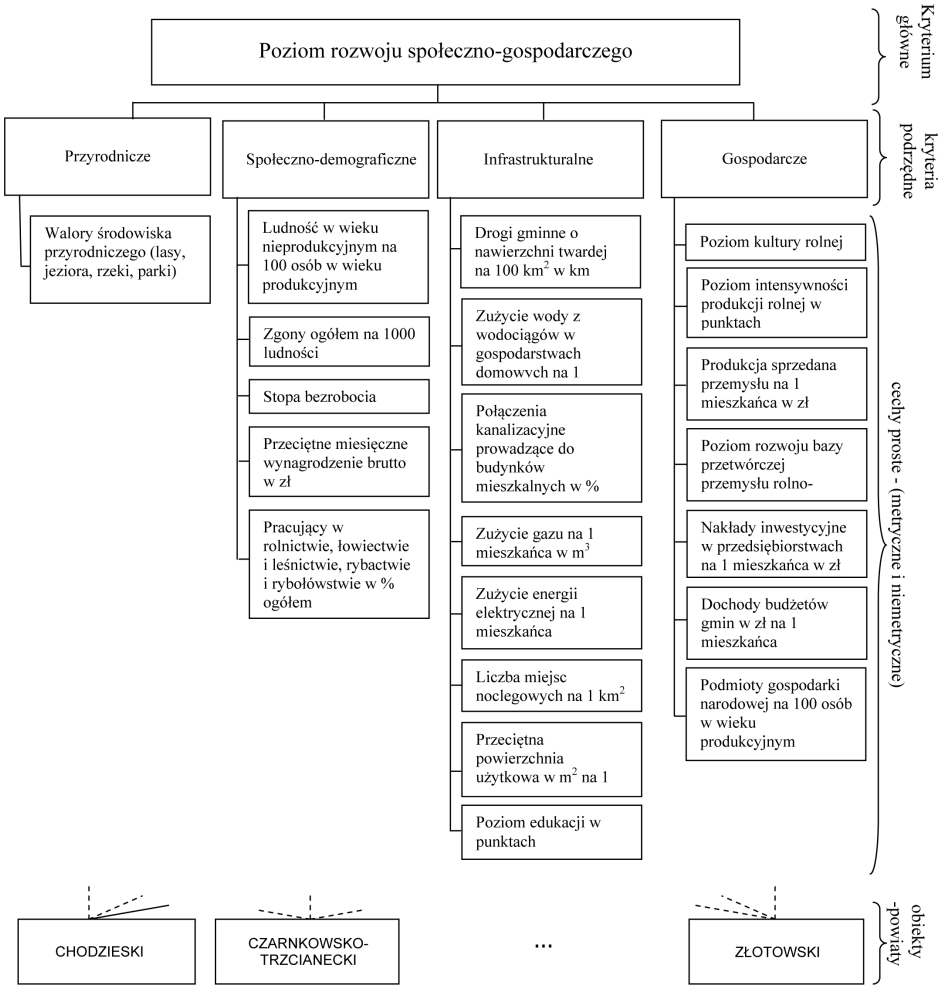
3. ZASTOSOWANIE ROZMYTYCH METOD AHP I TOPSIS W BADANIU POZIOMU ROZWOJU SPOŁECZNO-GOSPODARCZEGO POWIATÓW

Proponowaną w pracy wielokryterialną rozmytą metodę porządkowania liniowego zastosowano do oceny poziomu rozwoju społeczno-gospodarczego powiatów ziemskich województwa wielkopolskiego.

W pierwszym etapie utworzono strukturę hierarchiczną problemu oceny poziomu rozwoju społeczno-gospodarczego powiatów (rys. 3). Wśród wyróżnionych cech znajdują się cechy zarówno o charakterze metrycznym, jak i niemetrycznym (np. poziom kultury rolnej).

Stosując rozmytą metodę analitycznego procesu hierarchicznego według propozycji podanej przez Changa [2] obliczono najpierw współczynniki wagowe cech (etap 2). Przedstawimy przykład obliczenia współczynników wagowych dla cech w ramach kryterium demograficzno-społecznego. Idea tej metody polega na tym, że eksperci oceniają ważność poszczególnych cech w ramach kryterium, a ich oceny indywidualne zostają uśrednione (np. za pomocą średniej geometrycznej lub mediany). Porównując ważność cechy *ludność w wieku nieprodukcyjnym na 100 osób w wieku produkcyjnym* z cechą *zgony ogółem na 1000 ludności* ustalono, że pierwsza cecha jest mocno preferowana nad drugą i w związku z tym przypisano jej przewagę ważności $\tilde{5}$, co odpowiada trójkątnej liczbie rozmytej (3, 5, 7) (krok 1, tab. 3). Jednocześnie oznacza to, że *zgony ogółem na 1000 ludności* w porównaniu z liczbą *ludności w wieku nieprodukcyjnym na 100 osób w wieku produkcyjnym* otrzymają wagę $\frac{\tilde{1}}{\tilde{5}} = \frac{(1,\ 1,\ 1)}{(3,\ 5,\ 7)} = \left(\frac{1}{7}; \frac{1}{5}; \frac{1}{3}\right)$ (tab. 3).

W kroku 2 liczby rozmyte sumuje się dla każdej cechy oddzielnie. Dla cechy *ludność w wieku nieprodukcyjnym na 100 osób w wieku produkcyjnym* ($j = 2$, $k = 1$) obliczenia przebiegają w następujący sposób (wzór 2, *Dodatek*):



Rysunek 3. Struktura hierarchiczna problemu oceny poziomu rozwoju społeczno-gospodarczego powiatów województwa wielkopolskiego

Źródło: Opracowanie własne.

Tabela 3

Przykład obliczenia współczynników wagowych dla cech w ramach kryterium demograficzno-społecznego (metoda Changa [2], kroki 1-2)

Cechy		Ludność w wieku nieprodukcyjnym na 100 osób w wieku produkcyjnym			Zgony ogółem na 1000 ludności			Stopa bezrobocia			Przeciętne miesięczne wynagrodzenie brutto w zł			Pracujący w rolnictwie, łowiectwie i leśnictwie; rybactwo i rybołówstwie w % ogółem			$\sum_{g=1}^{5} l_{2kg}$	$\sum_{g=1}^{5} m_{2kg}$	$\sum_{g=1}^{5} u_{2kg}$	Znormalizowane wartości $\bar{Q}_{2k}$		
	g	1			2			3			4			5								
	k	l_{2k1}	m_{2k2}	u_{2k1}	l_{2k2}	m_{2k2}	u_{2k2}	l_{k3}	m_{2k3}	u_{2k3}	l_{2k4}	m_{2k4}	u_{2k4}	l_{2k5}	m_{2k5}	u_{2k5}				$\frac{\sum_{g=1}^{5} l_{2kg}}{\sum_{k=1}^{5}\sum_{g=1}^{5} u_{2kg}}$	$\frac{\sum_{g=1}^{5} m_{2kg}}{\sum_{k=1}^{5}\sum_{g=1}^{5} m_{2kg}}$	$\frac{\sum_{g=1}^{5} u_{2kg}}{\sum_{k=1}^{5}\sum_{g=1}^{5} l_{2kg}}$
Ludność w wieku nieprodukcyjnym na 100 osób w wieku produkcyjnym	1	1,00	1,00	1,00	3,00	5,00	7,00	$\frac{1}{7}$	$\frac{1}{5}$	$\frac{1}{3}$	1,00	3,00	5,00	$\frac{1}{7}$	$\frac{1}{5}$	$\frac{1}{3}$	5,286	9,400	13,667	0,068	0,170	0,396
Zgony ogółem na 1000 ludności	2	$\frac{1}{7}$	$\frac{1}{5}$	$\frac{1}{3}$	1,00	1,00	1,00	0,11	$\frac{1}{7}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{3}$	1,00	$\frac{1}{7}$	$\frac{1}{5}$	$\frac{1}{3}$	1,597	1,876	2,867	0,020	0,034	0,083
Stopa bezrobocia	3	3,00	5,00	7,00	5,00	7,00	9,00	1,00	1,00	1,00	5,00	7,00	9,00	1,00	3,00	5,00	15,000	23,000	31,000	0,192	0,416	0,898
Przeciętne miesięczne wynagrodzenie brutto w zł	4	$\frac{1}{5}$	$\frac{1}{3}$	1,00	1,00	3,00	5,00	$\frac{1}{9}$	$\frac{1}{7}$	$\frac{1}{5}$	1,00	1,00	1,00	$\frac{1}{7}$	$\frac{1}{5}$	$\frac{1}{3}$	2,454	4,676	7,533	0,031	0,085	0,218
Pracujący w rolnictwie, łowiectwie i leśnictwie; rybactwo i rybołówstwie w % ogółem	5	3,00	5,00	7,00	3,00	5,00	7,00	$\frac{1}{5}$	$\frac{1}{3}$	1,00	3,00	5,00	7,00	1,00	1,00	1,00	10,200	16,333	23,000	0,131	0,295	0,666
																Σ	34,537	55,286	78,067	0,442	1,000	2,260

Źródło: Obliczenia własne.

$$\sum_{g=1}^{5} l_{21g} = 1 + 3 + \frac{1}{7} + 1 + \frac{1}{7} = 5,286,$$

$$\sum_{g=1}^{5} m_{21g} = 1 + 5 + \frac{1}{5} + 3 + \frac{1}{5} = 9,400,$$

$$\sum_{g=1}^{5} u_{21g} = 1 + 7 + \frac{1}{3} + 5 + \frac{1}{3} = 13,667$$

$$\sum_{k=1}^{5}\sum_{g=1}^{5} l_{2kg} = 5{,}286 + 1{,}597 + 15{,}000 + 2{,}454 + 10{,}200 = 34{,}537,$$

$$\sum_{k=1}^{5}\sum_{g=1}^{5} m_{2kg} = 9{,}400 + 1{,}876 + 23{,}000 + 4{,}676 + 16{,}333 = 55{,}286,$$

$$\sum_{k=1}^{5}\sum_{g=1}^{5} u_{2kg} = 13{,}667 + 2{,}867 + 31{,}000 + 7{,}533 + 23{,}000 = 78{,}067.$$

Wtedy uzyskuje się:
$\tilde{Q}_{21} = (l_{21}, m_{21}, u_{21}) = (5,286;\ 9,400;\ 13,667) \otimes (34,537;\ 55,286;\ 78,067)^{-1} =$
$= \left(\frac{5,286}{78,067}; \frac{9,400}{55,286}; \frac{13,667}{34,537}\right) = (0{,}068;\ 0{,}170;\ 0{,}396).$

W kroku trzecim oblicza się na podstawie wzoru (3) stopnie możliwości dla $V\left(\tilde{Q}_{jk} \geqslant \tilde{Q}_{jg}\right)$. W przypadku kryterium demograficzno-społecznego ($j = 2$) należy uwzględnić pięć cech (g, $k = 1, \ldots, 5$). Porównując $\tilde{Q}_{21}$ z $\tilde{Q}_{2g}$, dla g= 2, 3, 4, 5, otrzymujemy kolejno:

1) $\tilde{Q}_{21}$= (0,068; 0,170; 0,396) i $\tilde{Q}_{22}$= (0,020; 0,034; 0,083), stąd $m_{21} \geqslant m_{22}$ i $V\left(\tilde{Q}_{21} \geqslant \tilde{Q}_{22}\right) = 1$,

2) $\tilde{Q}_{21}$= (0,068; 0,170; 0,396) i $\tilde{Q}_{23}$= (0,192; 0,416; 0,898), zachodzi trzecia ewentualność we wzorze (3), a zatem

$$V\left(\tilde{Q}_{21} \geqslant \tilde{Q}_{23}\right) = \frac{l_{23} - u_{21}}{(m_{21} - u_{21}) - (m_{23} - l_{23})} =$$

$$\frac{0,192 - 0,396}{(0,170 - 0,396) - (0,416 - 0,192)} = 0,453,$$

3) $\tilde{Q}_{21}$= (0,068; 0,170; 0,396) i $\tilde{Q}_{24}$= (0,031; 0,085; 0,218), stąd $m_{21} \geqslant m_{24}$ i $V\left(\tilde{Q}_{21} \geqslant \tilde{Q}_{24}\right) = 1$,

4) $\tilde{Q}_{21}$ = (0,068; 0,170; 0,396) i $\tilde{Q}_{25}$ = (0,131; 0,295; 0,666), zachodzi trzecia ewentualność we wzorze (3), a zatem

$$V\left(\tilde{Q}_{21} \geqslant \tilde{Q}_{25}\right) = \frac{l_{25} - u_{21}}{(m_{21} - u_{21}) - (m_{25} - l_{25})} =$$

$$\frac{0,131 - 0,396}{(0,170 - 0,396) - (0,295 - 0,131)} = 0,679.$$

Zgodnie ze wzorem (4) wartość minimalna stopnia możliwości (krok 4) wyniesie: $V\left(\tilde{Q}_{21} \geqslant \tilde{Q}_{2g} | g = 1, ..., 5 \; ; g \neq 1\right) = \min(1;\ 0,453;\ 1;\ 0,679) = 0,453$.

W kroku piątym obliczamy współczynniki wagowe, najpierw jako wagi lokalne (tab. 4). Dla omawianej cechy (k = 1, j = 2) współczynnik ten wynosi (wzór 5):

$$w_{21}^{(l)} = \frac{V\left(\tilde{Q}_{21} \geqslant \tilde{Q}_{2g} | g = 2, ..., 5\right)}{\sum_{h=1}^{5} V\left(\tilde{Q}_{2h} \geqslant \tilde{Q}_{2g} | g = 2, ..., 5\right)} = 0,453/2,323 = 0,195.$$

Aby otrzymać globalny współczynnik wagowy należy pomnożyć $w_{21}^{(l)} = 0,195$ przez współczynnik wagowy dla kryterium demograficzno-społecznego, który wynosi w_2= 0,262. Wtedy uzyskuje się $w_{21}^{=} w_{21}^{(l)} \cdot w_1^{=} 0,195 \cdot 0,262 = 0,051$ (krok 6).

Analogicznie według kroków (1)-(3) obliczono globalne współczynniki wagowe dla pozostałych cech (tab. 5). Zauważmy, że z badanego ich zbioru zostały wyeliminowane trzy cechy: *zgony ogółem na 1000 ludności* (kryterium społeczno-demograficzne), *liczba miejsc noclegowych na 1 km*^{2} (kryterium infrastrukturalne), *poziom kultury rolnej* (kryterium gospodarcze). Dla wymienionych cech współczynniki wagowe przyjęły wartość zero.

Otrzymane wektory współczynników wagowych dla cech stanowiły podstawę do zastosowania rozmytej metody TOPSIS. W etapie trzecim, w wyniku zamiany wartości cech na trójkątne liczby rozmyte otrzymuje się rozmytą macierz danych. Jej fragment zamieszczono w tab. 6. W kolejnym etapie stosując wzory (3)-(8) dokonuje się normalizacji liczb rozmytych (tab. 7). Zastosowano przekształcenia ilorazowe wykonywane na liczbach rozmytych. Znormalizowane liczby rozmyte zostały pomnożone przez współczynniki wagowe zgodnie ze wzorem (9). Na ich podstawie zostały wyznaczone wartości rozmytego wzorca i antywzorca rozwoju według formuł (10 - 11):

$$\tilde{A}^{+} = ((0,044,\ 0,055,\ 0,055);\ (0,052,\ 0,052,\ 0,052),...,(0,140,\ 0,140,\ 0,140))$$

$$\tilde{A}^{-} = ((0,00;\ 0,\ 000;\ 0,011),\ (0,042,\ 0,042,\ 0,042), ..., (0,062,\ 0,062,\ 0,062))$$

Tabela 4

Przykład obliczania współczynników wagowych dla cech w ramach kryterium demograficzno-społecznego (metoda Changa [2], krok 3-5)

Numer cechy k	$\tilde{Q}_{2k}$			Numer cechy g	$\tilde{Q}_{2g}$			$V\left(\tilde{Q}_{2k} \geqslant \tilde{Q}_{2g}\right)$	k	$\min V(\tilde{Q}_{2k} \geqslant \tilde{Q}_{2g})$	$w_{2k}^{(l)}$	w_{2k}	Cechy w ramach kryterium demograficzno-społecznego (j=2)
	l_{2k}	m_{2k}	u_{2k}		l_{2g}	m_{2g}	u_{2g}						
1	0,068	0,170	0,396	2	0,020	0,034	0,083	1,000					
1	0,068	0,170	0,396	3	0,192	0,416	0,898	0,453	1	0,453	0,195	0,051	Ludność w wieku nieprodukcyjnym na 100 osób w wieku produkcyjnym
1	0,068	0,170	0,396	4	0,031	0,085	0,218	1,000					
1	0,068	0,170	0,396	5	0,131	0,295	0,666	0,679					
2	0,020	0,034	0,083	1	0,068	0,170	0,396	0,101					
2	0,020	0,034	0,083	3	0,192	0,416	0,898	0,000	2	0,000	0,000	0,000	**Zgony ogółem na 1000 ludności**
2	0,020	0,034	0,083	4	0,031	0,085	0,218	0,505					
2	0,020	0,034	0,083	5	0,131	0,295	0,666	0,000					
3	0,192	0,416	0,898	1	0,068	0,170	0,396	1,000					
3	0,192	0,416	0,898	2	0,020	0,034	0,083	1,000	3	1,000	0,431	0,113	Stopa bezrobocia (w %)
3	0,192	0,416	0,898	4	0,031	0,085	0,218	1,000					
3	0,192	0,416	0,898	5	0,131	0,295	0,666	1,000					
4	0,031	0,085	0,218	1	0,068	0,170	0,396	0,638					
4	0,031	0,085	0,218	2	0,020	0,034	0,083	1,000	4	0,073	0,031	0,008	Przeciętne miesięczne wynagrodzenie brutto w zł
4	0,031	0,085	0,218	3	0,192	0,416	0,898	0,073					
4	0,031	0,085	0,218	5	0,131	0,295	0,666	0,293					
5	0,131	0,295	0,666	1	0,068	0,170	0,396	1,000					
5	0,131	0,295	0,666	2	0,020	0,034	0,083	1,000	5	0,797	0,343	0,090	Pracujący w rolnictwie, łowiectwie i leśnictwie, rybactwie i rybołówstwie w % ogółem
5	0,131	0,295	0,666	3	0,192	0,416	0,898	0,797					
5	0,131	0,295	0,666	4	0,030	0,080	0,220	1,000					
									Σ	2,323	1,000	0,262	×

Źródło: Obliczenia własne na podstawie danych tab. 3.

Tabela 5

Wagi kryteriów i cech opisujących sytuację społeczno-gospodarczą powiatów w województwie wielkopolskim (uzyskane metodą Changa [2])

Kryteria i cechy	Wagi	
	lokalne $w_{jk}^{(l)}$	globalne w_j i w_{jk}
Przyrodnicze (w_1)	1,000	0,055
Walory środowiska przyrodniczego (lasy, jeziora, rzeki, parki) (w punktach)	1,000	0,055
Demograficzno-społeczne (w_2)	1,000	0,262
Ludność w wieku nieprodukcyjnym na 100 osób w wieku produkcyjnym	0,195	0,051
Zgony ogółem na 1000 ludności	**0,000**	**0,000**
Stopa bezrobocia (%)	0,431	0,113
Przeciętne miesięczne wynagrodzenie brutto w zł	0,031	0,008
Pracujący w rolnictwie, łowiectwie i leśnictwie, rybactwie i rybołówstwie w % ogółem	0,343	0,090
Infrastrukturalne (w_3)	0,999	0,118
Drogi gminne o nawierzchni twardej na 100 km^2 w km	0,252	0,030
Zużycie wody z wodociągów w gospodarstwach domowych na 1 mieszkańca	0,073	0,009
Połączenia kanalizacyjne prowadzące do budynków mieszkalnych w % ogółu budynków	0,182	0,021
Zużycie gazu na 1 mieszkańca w m^3	0,075	0,009
Zużycie energii elektrycznej (kWh) na 1 mieszkańca	0,066	0,008
Liczba miejsc noclegowych na 1 km^2	**0,000**	**0,000**
Przeciętna powierzchnia użytkowa w m^2 na 1 osobę	0,086	0,010
Poziom edukacji (w punktach)	0,265	0,031
Gospodarcze (w_4)	1,000	0,564
Poziom kultury rolnej (w punktach)	**0,000**	**0,000**
Poziom intensywności produkcji rolnej (w punktach)	0,013	0,007
Produkcja sprzedana przemysłu na 1 mieszkańca w zł	0,257	0,145
Poziom rozwoju bazy przetwórczej (w punktach)	0,087	0,049
Nakłady inwestycyjne w przedsiębiorstwach na 1 mieszkańca w zł	0,225	0,127
Dochody budżetów gmin w zł na 1 mieszkańca	0,170	0,096
Podmioty gospodarki narodowej na 100 osób w wieku produkcyjnym	0,248	0,140

Źródło: Obliczenia własne na podstawie [1, 11, 15].

T a b e l a 6

Wartości cech oraz odpowiadające im trójkątne liczby rozmyte (fragment macierzy danych)

Nr	Powiaty	Cechy			
		Walory środowiska przyrodniczego (lasy, jeziora, rzeki, parki)	Ludność w wieku nieprodukcyjnym na 100 osób w wieku produkcyjnym	...	Podmioty gospodarki narodowej na 100 osób w wieku produkcyjnym
1	Chodzieski	BW[a)]	57,2		13,9
		(80; 100; 100)[b)]	(57,2; 57,2; 57,2)	...	(13,9; 13,9; 13,9)
...	...	...	...	...	...
31	Złotowski	Ś	61,5		9,5
		(40; 50; 60)	(61,5; 61,5; 61,5)	...	(9,5; 9,5; 9,5)
	max	(80; 100; 100)	(66,5; 66,5; 66,5)		(18,74; 18,74; 18,74)
	min	(0, 0, 20)	(54,3; 54,3; 54,3)		(8,36; 8,36; 8,36)
	Współczynniki wagowe	0,055	0,051		0,140

[a)] Wartości cechy określane jako poziomy zmiennej lingwistycznej: bardzo wysoki (BW), wysoki (W), średni (Ś), niski (N), bardzo niski (BN) albo liczby rzeczywiste.
[b)] Trójkątna liczba rozmyta.
Źródło: Obliczenia własne na podstawie wyników badania ankietowego przeprowadzonego w starostwach powiatowych województwa wielkopolskiego (2001) [16], [11, 15].

oraz obliczone odległości powiatów od wzorca i antywzorca rozwoju (etap 3). Na przykład odległość powiatu chodzieskiego od wzorca rozwoju wyniesie (wzory 12-13):

$$d_1^+ = \sum_{k=1}^{18} d\left(\tilde{r}_{1k}; \tilde{r}_k^+\right) = \sqrt{\frac{1}{3}\left[(0,044-0,044)^2 + (0,055-0,055)^2 + (0,055-0,055)^2\right]} +$$

$$+ ... + \sqrt{\frac{1}{3}\left[(0,104-0,140)^2 + (0,104-0,140)^2 + (0,104-0,140)^2\right]} = 0,453 \quad ,$$

oraz od antywzorca:

$$d_1^- = \sum_{k=1}^{18} d\left(\tilde{r}_{1k}; \tilde{r}_k^-\right) = \sqrt{\frac{1}{3}\left[(0,044-0,000)^2 + (0,055-0,000)^2 + (0,055-0,011)^2\right]} +$$

$$+ ... + \sqrt{\frac{1}{3}\left[(0,104-0,062)^2 + (0,104-0,062)^2 + (0,104-0,062)^2\right]} = 0,221$$

Obliczone odległości posłużyły do wyznaczenia wartości cechy syntetycznej według wzoru (14). Wynik obliczeń dla powiatu chodzieskiego jest następujący:

$$S_1 = \frac{d_1^-}{d_1^+ + d_1^-} = \frac{0,221}{0,453 + 0,221} = \frac{0,221}{0,674} = 0,328$$

Tabela 7

Znormalizowane wartości rozmyte wybranych cech przemnożone przez współczynniki wagowe (fragment macierzy)

Nr	Powiaty	Cechy			
		Walory środowiska przyrodniczego (lasy, jeziora, rzeki, parki)	Ludność w wieku nieprodukcyjnym na 100 osób w wieku produkcyjnym	...	Podmioty gospodarki narodowej na 100 osób w wieku produkcyjnym
1	Chodzieski	0,055·(80/100; 100/100; 100/100)= = 0,055· (0,8; 1; 1) = = (0,044; 0,055; 0,055)	0,051 (54,3/57,2; 54,3/57,2; 54,3/57,2)= 0,051 (0,95; 0,95; 0,95) = = (0,050; 0,050; 0,050)	...	0,140· (13,9/18,74; 13,9/18,74; 13,9/18,74) = = 0,140· (0,74;0,74; 0,74) = = (0,104; 0,104; 0,104)
...	...				
31	Złotowski	0,055· (40/100; 50/100; 60/100)= = 0,055·(0,40; 0,50; 0,60)= = (0,022; 0,028; 0,033)	0,051 (54,3/61,5; 54,3/61,5; 54,3/61,5) = = 0,051 (0,88; 0,88; 0,88) = = (0,046; 0,046; 0,046)	...	0,140· (9,5/18,74; 9,5/18,74; 9,5/18,74) = = 0,140· (0,51; 0,51; 0,51) = = (0,071; 0,071; 0,071)
	wzorzec	(0,044; 0,055; 0,055)	(0,052, 0,052, 0,052)		(0,140, 0, 140, 0, 140)
	antywzorzec	(0,000; 0,000; 0,011)	(0,042; 0,042; 0,042)		(0,062; 0,062; 0,062)

Źródło: Obliczenia własne na podstawie tab. 5 i 6.

Tabela 8

Uporządkowanie liniowe powiatów województwa wielkopolskiego według poziomu rozwoju społeczno-gospodarczego

Lp.	Powiaty	Wartości cechy syntetycznej S_i/ metody		Klasa i poziom rozwoju[c)]
		rozmyta AHP i TOPSIS[a)]	klasyczna Hellwiga[b)]	
1	Poznański	0,712	0,394	I wysoki
2	Gostyński	0,508	0,327	II
3	Szamotulski	0,493	0,207	średni-wyższy
4	Kościański	0,475	0,283	
5	Wolsztyński	0,472	0,355	
6	Pilski	0,428	0,260	
7	Leszczyński	0,424	0,234	
8	Grodziski	0,410	0,261	III
9	Krotoszyński	0,384	0,205	
10	Międzychodzki	0,367	0,167	
11	Kępiński	0,364	0,169	średni-niższy
12	Obornicki	0,359	0,109	
13	Średzki	0,359	0,213	
14	Rawicki	0,356	0,203	
15	Nowotomyski	0,355	0,141	
16	Śremski	0,332	0,160	
17	Chodzieski	0,328	0,187	
18	Ostrzeszowski	0,328	0,193	
19	Ostrowski	0,325	0,169	
20	Kaliski	0,317	0,069	IV
21	Gnieźnieński	0,305	0,245	niski-wyższy
22	Wrzesiński	0,299	0,125	
23	Jarociński	0,286	0,112	
24	Czarnkowsko-trzcianecki	0,273	0,050	
25	Wągrowiecki	0,262	0,102	
26	Koniński	0,260	0,166	
27	Turecki	0,251	0,103	
28	Kolski	0,244	0,053	
29	Pleszewski	0,223	0,106	V
30	Słupecki	0,200	0,080	niski
31	Złotowski	0,188	0,029	

[a)] Uporządkowanie liniowe według wartości syntetycznego miernika rozwoju uzyskanego rozmytymi metodami AHP i TOPSIS.

[b)] Bez uwzględnienia cech porządkowych i wag dla cech.

[c)] Podziału na klasy dokonano za pomocą średniej arytmetycznej i odchylenia standardowego obliczonych z wartości syntetycznego miernika rozwoju.

Źródło: Obliczenia własne na podstawie wyników badania ankietowego przeprowadzonego w starostwach powiatowych województwa wielkopolskiego (2001) [16] oraz [1, 11, 15].

Tabela 8 pokazuje uporządkowanie liniowe badanych powiatów według nierosnących wartości rzeczywistych cechy syntetycznej (etap 4). Dla porównania obliczono wartości cechy syntetycznej klasyczną metodą Hellwiga. Rozmyta metoda TOPSIS – poprzez uwzględnienie cech metrycznych i niemetrycznych – porządkowych, ich wag oraz odniesienia wartości cech do wzorca i antywzorca rozwoju – dostarczyła, w porównaniu z metodą Hellwiga [5], większego zakresu zmienności wartości syntetycznego miernika rozwoju (tab. 8). Rozstęp pomiędzy maksymalną a minimalną wartością syntetycznego miernika rozwoju uzyskany rozmytą metodą TOPSIS wynosi 0,524, a w przypadku metody Hellwiga 0,365. Ponadto wartości miernika uzyskane metodą Hellwiga wskazują na niski bądź bardzo niski poziom rozwoju powiatów, co budzi pewne wątpliwości.

Na podstawie uporządkowanych wartości cechy syntetycznej uzyskanych rozmytą metodą TOPSIS wyodrębniono pięć typów rozwojowych powiatów (tab. 8). Pierwszy typ utworzył powiat poznański, najlepiej rozwinięty pod względem społeczno-gospodarczym. Istotny wpływ na rozwój tego powiatu ma oddziaływanie aglomeracji miejskiej Poznania. Drugi typ obejmuje trzy powiaty: kościański, gostyński i szamotulski. Są to tereny charakteryzujące się średnim-wyższym poziomem rozwoju. Trzeci typ tworzy jedenaście powiatów głównie z południowo-zachodniej części województwa. Powiaty te cechują się średnim-niższym poziomem rozwoju. Czwarty typ występuje na obszarze dwunastu powiatów głównie z północnej i południowej części województwa. Są to tereny o niskim-wyższym poziomie rozwoju. Ostatni piąty typ to tereny o niskim poziomie rozwoju społeczno-gospodarczego. Ten typ występuje w czterech powiatach położonych peryferyjnie, we wschodniej i północnej części województwa.

4. PODSUMOWANIE

Na podstawie przeprowadzonych badań i analiz można sformułować następujące stwierdzenia i wnioski:

1. Zaproponowana metoda porządkowania liniowego obiektów wykorzystująca procedury rozmyte AHP i TOPSIS jest przydatna w procedurze tworzenia cechy syntetycznej. Jej zasadnicze zalety w porównaniu z metodą klasyczną Hellwiga można upatrywać w możliwości uwzględnienia w procesie budowy miernika syntetycznego zarówno cech o charakterze metrycznym, jak i niemetrycznym (porządkowym), systemu wag dla cech ustalonych metodami eksperckimi, a także możliwości odniesienia wartości cech zarówno do wzorca jak i antywzorca rozwoju.
2. Za pomocą rozmytego analitycznego procesu hierarchicznego poszczególnym kryteriom, jak i cechom można przyporządkować zróżnicowane współczynniki wagowe, a także wyeliminować cechy o najmniejszym znaczeniu w sensie merytorycznym (w opinii ekspertów). W prezentowanym przykładzie analizowanych było 21 cech opisujących powiaty. Z tego zbioru wyeliminowane zostały cechy: zgony ogółem na 1000 ludności (kryterium społeczno-demograficzne), liczba miejsc noc-

legowych na 1 km^2 (kryterium infrastrukturalne), poziom kultury rolnej (kryterium gospodarcze).

3. Stosując zaproponowaną metodę uzyskano znacznie większy zakres zmienności wartości syntetycznego miernika rozwoju w porównaniu z klasyczną metodą Hellwiga (odpowiednio 0,524 i 0,365). Według metody Hellwiga poziom rozwoju wszystkich powiatów trzeba by uznać za niski bądź bardzo niski (wartości miernika $S_i \in (0,0;\ 0,4)$, co oczywiście nie odpowiada rzeczywistości. Przy tym samym zakresie zmienności mierników TOPSIS i Hellwiga (0, 1) ich wartość na przykład dla powiatu poznańskiego wyniosła odpowiednio 0,712 i 0,394. Oznacza to, że poziom rozwoju społeczno-gospodarczego powiatu poznańskiego według miernika uzyskanego rozmytymi metodami AHP i TOPSIS został oceniony jako „wysoki", podczas gdy z wykorzystaniem metody Hellwiga jako „średni-niższy".

Uniwersytet Przyrodniczy w Poznaniu

LITERATURA

[1] *Bank Danych Regionalnych* [2006], GUS, www.stat.gov.pl/bdr_s/app/strona.indeks.
[2] Chang D.-Y. [1996], *Application of the Extent Analysis Method on fuzzy AHP*, „European Journal of Operational Research", 95 (2), s. 649-655.
[3] Chang Y. - H, Yeh, C.-H. [2004], *A new airline safety index*. „Transportation Research Part B" 38, s. 369-383.
[4] Chen C.-T. [2000], *Extensions of the TOPSIS for group decision-making under fuzzy environment.* „Fuzzy Sets and Systems" 114 (1), s. 1-9.
[5] Hellwig Z. [1968]: Zastosowania metody taksonomicznej do typologicznego podzialu krajów ze wzgledu na poziom ich rozwoju i strukture wykwalifikowanych kadr. *Przeglad Statystyczny*, z. 4, str. 307-327.
[6] Kukuła K. [2000], *Metoda unitaryzacji zerowej*. Wyd. PWN, Warszawa, s. 64.
[7] Łuczak A., Wysocki F. [2010]: Wykorzystanie rozmytych metod AHP i TOPSIS do porządkowania liniowego obiektów. *Taksonomia* 17, Klasyfikacja i analiza danych. Teoria i zastosowania, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław 2010, s. 334-343.
[8] Malina A., Zeliaś A. [1997]: *Taksonomiczna analiza przestrzennego zróżnicowania jakości życia ludności w Polsce w 1994 r. Przegląd Statystyczny* 44 (1), s. 11-27.
[9] Nowak E. [1990]: *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych.* PWE, Warszawa.
[10] Saaty T. L., [1980]: *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*, MacGraw- Hill, New York International Book Company.
[11] *Strategia rozwoju rolnictwa i obszarów wiejskich w Wielkopolsce,* [2000], pod red. W. Poczta i F. Wysocki, Sejmik Województwa Wielkopolskiego. Poznań.
[12] Walesiak M. [1993]: *Statystyczna analiza wielowymiarowa w badaniach marketingowych*, Prace Naukowe AE we Wrocławiu, Nr 654, Seria: Monografie i opracowania nr 101.
[13] Wang J.-W., Cheng C.-H., Kun-Cheng H., [2009], *Fuzzy hierarchical TOPSIS for supplier selection*. „Applied Soft Computing", 9, s. 377–386
[14] Wang Y.-M., Luo Y., Hua Z. [2008], *On the extent analysis method for fuzzy AHP and its applications*. „European Journal of Operational Research", 186, s. 735-747.
[15] *Ważniejsze dane o powiatach i gminach województwa wielkopolskiego 2004,* [2004], WUS, Poznań.

[16] Wyniki badania ankietowego przeprowadzonego w starostwach powiatowych województwa wielkopolskiego [2001].
[17] Wysocki F. [2010]: *Metody taksonomiczne w rozpoznawaniu typów ekonomicznych rolnictwa i obszarów wiejskich*. Wydawnictwo Uniwersytetu Przyrodniczego w Poznaniu, Poznań.

PORZADKOWANIE LINIOWE OBIEKTÓW Z WYKORZYSTANIEM ROZMYTYCH METOD AHP I TOPSIS

Streszczenie

Celem pracy było przedstawienie możliwości zastosowania rozmytej wielokryterialnej metody porządkowania liniowego do konstrukcji cechy syntetycznej. Metoda polega na wykorzystaniu dwóch komplementarnych rozmytych metod: analitycznego procesu hierarchicznego do ustalenia wag cech prostych oraz rozmytej metody TOPSIS przy rangowaniu obiektów. Zaproponowana procedura została zilustrowana przykładem dotyczącym oceny poziomu rozwoju społeczno-gospodarczego powiatów województwa wielkopolskiego.

Słowa kluczowe: rozmyty AHP, rozmyta metoda TOPSIS, porządkowanie liniowe obiektów

LINEAR ORDERING OF OBJECTS FROM APPLICATION OF FUZZY AHP AND TOPSIS

Summary

The aim of this paper was to investigate the applicability of the fuzzy multi-criteria linear ordering method to the construction of synthetic characteristics. Proposed approach base on two fuzzy methods: analytical hierarchy process (to calculate weight of characteristics and eliminated of unimportance characteristics) and method TOPSIS (to ranking of objects). The proposed procedure was employed to assess the socio-economic development of rural Wielkopolska seen as a collection of counties.

Key words: fuzzy AHP, fuzzy TOPSIS, linear ordering of objects

Dodatek – działania na liczbach rozmytych

dodawanie:

$$\tilde{x}_1 \oplus \tilde{x}_2 = (a_1, b_1, c_1) \oplus (a_2, b_2, c_2) = (a_1 + a_2, b_1 + b_2, c_1 + c_2)$$

odejmowanie:

$$\tilde{x}_1 - \tilde{x}_2 = (a_1, b_1, c_1) - (a_2, b_2, c_2) = (a_1 - c_2, b_1 - b_2, c_1 - a_2)$$

mnożenie:

$$\tilde{x}_1 \otimes \tilde{x}_2 = (a_1, b_1, c_1) \otimes (a_2, b_2, c_2) \cong (a_1 a_2, b_1 b_2, c_1 c_2) \quad \tilde{x}_1, \tilde{x}_2 > 0$$

$$\tilde{x}_1 \otimes \tilde{x}_2 = (a_1, b_1, c_1) \otimes (a_2, b_2, c_2) \cong (c_1 a_2, b_1 b_2, a_1 c_2) \quad \tilde{x}_1 > 0, \tilde{x}_2 < 0$$

$$\tilde{x}_1 \otimes \tilde{x}_2 = (a_1, b_1, c_1) \otimes (a_2, b_2, c_2) \cong (c_1 c_2, b_1 b_2, a_1 a_2) \quad \tilde{x}_1, \tilde{x}_2 < 0$$

$$\tilde{x}_1 \otimes \tilde{x}_2 = (a_1, b_1, c_1) \otimes (a_2, b_2, c_2) \cong (a_1 c_2, b_1 b_2, c_1 a_2) \quad \tilde{x}_1 < 0, \tilde{x}_2 > 0$$

dzielenie:

$$\frac{\tilde{x}_1}{\tilde{x}_2} = (a_1, b_1, c_1) \ / \ (a_2, b_2, c_2) \cong \left(\frac{a_1}{c_2}, \frac{b_1}{b_2}, \frac{c_1}{a_2} \right)$$

liczba przeciwna do liczby rozmytej:

$$-(a, b, c) = (-c, -b, -a)$$

dodawanie skalaru do liczby rozmytej:

$$k \oplus (a, b, c) = (k + a, k + b, k + c)$$

mnożenie skalaru przez liczbę rozmytą:

$$k \otimes (a, b, c) \cong (ka, kb, kc)$$

PRZEGLĄD STATYSTYCZNY
R. LVIII – ZESZYT 1-2 – 2011

BEATA JACKOWSKA

EFEKTY INTERAKCJI MIĘDZY ZMIENNYMI OBJAŚNIAJĄCYMI W MODELU LOGITOWYM W ANALIZIE ZRÓŻNICOWANIA RYZYKA ZGONU

1. WSTĘP

Modele regresji logistycznej (modele logitowe) wykorzystywane są do objaśniania zmiennych jakościowych w zależności od poziomu zmiennych egzogenicznych (jakościowych bądź ilościowych). Regresja logistyczna znajduje ważne zastosowanie m. in. w modelowaniu ryzyka znalezienia się jednostki badania w pewnym stanie. Jeżeli zmienna objaśniana przyjmuje dwa stany, tzn. mówi, czy badane zjawisko wystąpiło, czy też nie, to mamy do czynienia z modelem dwumianowym.[1]

Współcześnie modele logitowe znalazły powszechne zastosowanie w bankach do oceny ryzyka kredytowego oraz w przedsiębiorstwach do oceny lojalności klientów. Są także jednym z narzędzi wykorzystywanym przez aktuariuszy do oceny ryzyka ubezpieczeniowego oraz oceny szansy konwersji i retencji polis ubezpieczeniowych [7]. W ubezpieczeniach na życie model logitowy pozwala na oszacowanie prawdopodobieństwa śmierci w zależności od podstawowych cech demograficznych, takich jak płeć, wiek, miejsce zamieszkania [9], a w przypadku posiadania odpowiednio dużych baz danych dotyczących historii ubezpieczonych do modelu można włączyć informacje zbierane za pomocą ankiet medycznych dołączanych do wniosku ubezpieczeniowego.

W dziedzinie demografii na ogół poszukuje się parametrycznych (analitycznych) modeli ludzkiego procesu przeżycia (tzw. praw wymieralności[2]) jedynie w zależności od wieku budując oddzielnie modele dla kobiet i mężczyzn (ewentualnie dla osób mieszkających w miastach i na wsi). W tym celu wykorzystuje się regresję krzywoliniową. Jako modele analityczne współczynników zgonu najczęściej są stosowane funkcje: wykładnicze, potęgowe, wielomianowe, wielomianowo-wykładnicze, logistyczne (por. np. [11]). Dostęp do baz danych coraz lepszej jakości oraz rozwój metod numerycznych i oprogramowania komputerowego sprawiają, że modele te są coraz bardziej rozbudowane. W literaturze można odnaleźć wiele prób stworzenia demograficznych modeli analitycznych, lecz rzadziej spotkać można zastosowania uogólnionych modeli

[1] Model wielomianowy jest wykorzystywany, jeżeli zmienna zależna może przyjmować więcej niż dwa niezależne stany – mamy wówczas do czynienia z modelem ryzyka konkurencyjnego, tzn. wszystkie zdarzenia są wzajemnie niezależne i suma prawdopodobieństw ich wystąpienia wynosi 1.

[2] Do pierwszych znanych praw wymieralności należą m. in. prawo de Moivre (1725), prawo Gompertza (1825), prawo Makehama (1860), prawo Weibulla (1939). [11]

liniowych (GLM – *generalised linear models*), w tym regresji logistycznej, do analizy danych demograficznych [7], [9], [10]. Metody regresji logistycznej pozwalają na znalezienie statystycznie istotnych czynników ryzyka zgonu oraz zbadanie efektów interakcji między tymi czynnikami, a dodatkowym atutem modelu logitowego jest możliwość interpretacji jego parametrów.

O ile metody regresji logistycznej są szeroko opisane w literaturze i znajdują coraz więcej zastosowań, to mniej uwagi poświęca się modelowaniu interakcji [6]. Przyjmuje się, że efekt interakcji występuje, jeżeli wpływ zmiennej niezależnej na zmienną zależną zmienia się w zależności od wartości innej zmiennej niezależnej. Analiza efektów interakcji między zmiennymi objaśniającymi w regresji logistycznej została dokonana poprzez wprowadzenie do modelu iloczynu tych zmiennych. Szczególna uwaga została zwrócona na interpretację parametrów modelu, która zależy od sposobu kodowania zmiennych oraz uwzględnienia efektów interakcji między zmiennymi objaśniającymi.

Dopasowanie oraz zdolność predykcyjna modelu zależą od jakości danych oraz liczebności grup jednostek wyodrębnionych według wariantów analizowanych cech. W szczególności dla osób o zaawansowanym wieku ubezpieczyciele w Polsce nie posiadają wystarczająco dużo obserwacji, lecz mogą wykorzystać dane demograficzne gromadzone przez GUS. Estymacja modelu przeżycia dla osób starszych jest przede wszystkim niezbędna przy konstrukcji dobrowolnych ubezpieczeniowych produktów emerytalnych, jak dotąd słabo rozpowszechnionych w Polsce.

W niniejszym artykule dwumianowy model logitowy został wykorzystany do analizy ryzyka zgonu osób starszych (w wieku co najmniej 60 lat) w województwie pomorskim w 2009 roku w zależności od podstawowych cech demograficznych. Celem tej analizy była identyfikacja predyktorów ryzyka zgonu oraz odkrycie efektów interakcji między zmiennymi objaśniającymi.

2. POSTAĆ MODELU LOGITOWEGO

Dwumianowy model regresji logistycznej (model logitowy) wykorzystywany jest do objaśniania dychotomicznej zmiennej jakościowej Y w zależności od poziomu zmiennych egzogenicznych $X_1,\ X_2,\ \ldots,\ X_k$ (jakościowych bądź ilościowych). Zmienna objaśniana reprezentowana jest zwykle przez zmienną zero-jedynkową:

$$Y = \begin{cases} 1 & \text{zdarzenie wystąpiło} \\ 0 & \text{zdarzenie nie wystąpiło} \end{cases}. \tag{1}$$

Model logitowy jest szczególnym przypadkiem uogólnionego modelu liniowego [10]:

$$g(\mu) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k, \tag{2}$$

gdzie β_0 jest wyrazem wolnym; $\beta_1,\ \beta_2,\ \ldots,\ \beta_k$ są współczynnikami regresji; g jest funkcją wiążąca (*link function*) określającą związek średniej wartości zmiennej objaśnianej $\mu = E(Y|X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k)$ z liniową kombinacją predyktorów.

W modelu logitowym $\mu = p = P(Y = 1|X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k)$, a funkcja wiążąca nazywana logitem ma postać

$$g(p) = \text{logit}(p) = \ln\left(\frac{p}{1-p}\right). \tag{3}$$

Podsumowując, model logitowy można zapisać w następującej postaci

$$p = P(Y = 1|X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{\exp\left(\beta_0 + \sum_{i=1}^{k} \beta_i x_i\right)}{1 + \exp\left(\beta_0 + \sum_{i=1}^{k} \beta_i x_i\right)}. \tag{4}$$

Parametry modelu β_0, β_1, $\ldots$, β_k estymuje się najczęściej metodą największej wiarygodności maksymalizując logarytm funkcji wiarygodności względem parametrów modelu za pomocą iteracyjnych procedur numerycznych[3].

Jeżeli wśród zmiennych objaśniających znajdują się zmienne jakościowe, to wprowadza się je do modelu odpowiednio kodując (*dummy coding*, *indicator coding*) [1]. Zwykle, gdy zmienna ma m wariantów, to wprowadza się $m - 1$ zmiennych zero-jedynkowych (*dummy variables*). Grupa jednostek badania, dla której wartości wszystkich zmiennych objaśniających są równe zero nazywa się grupą referencyjną (*reference group*). Badacz kodując predyktory za pomocą zmiennych zero-jedynkowych ustala arbitralnie grupę referencyjną (np. wybierając grupę najliczniejszą, największego lub najmniejszego ryzyka), która stanie się grupą odniesienia przy interpretacji parametrów modelu.[4]

Zaletą modelu logitowego jest możliwość interpretacji parametrów e^{β_i}. W tym celu wykorzystuje się pojęcie szansy (*odds*) definiowanej jako iloraz prawdopodobieństwa wystąpienia zdarzenia oraz prawdopodobieństwa nie wystąpienia zdarzenia. W rozważanym modelu (4) szansę można wyrazić jako funkcję zmiennych objaśniających:

$$\frac{p}{1-p} = \gamma(x_1, x_2, \ldots, x_k) = \exp\left(\beta_0 + \sum_{i=1}^{k} \beta_i x_i\right). \tag{5}$$

W przypadku wyrazu wolnego, wartość e^{β_0} jest interpretowana jako szansa wystąpienia zjawiska w grupie referencyjnej.

[3] W pracy [3] przedstawiono metody estymacji w przypadku makrodanych, tzn. gdy „(...) w miejsce obserwacji 0-1 dla zmiennej Y mamy dane jedynie frakcje tych obserwacji w grupach jednostek, których indywidualne cechy nie są rozróżnialne" [3, s. 87].

[4] Dyskusję o innych sposobach kodowania można znaleźć w [5] i [8]. Sposób kodowania ma wpływ na wartość współczynników regresji oraz na interpretację parametrów modelu (zmienia się grupa referencyjna), natomiast nie wpływa na wartość prognozowanego prawdopodobieństwa.

Wpływ przyrostu wartości zmiennych niezależnych o Δx_i ($i = 1, 2, \ldots, k$) na szansę wystąpienia zjawiska można określić wyznaczając iloraz szans (*odds ratio*):

$$\psi(x_1, x_2, \ldots, x_k; \Delta x_1, \Delta x_2, \ldots, \Delta x_k) = \frac{\gamma(x_1 + \Delta x_1, x_2 + \Delta x_2, \ldots, x_k + \Delta x_k)}{\gamma(x_1, x_2, \ldots, x_k)} = \exp\left(\sum_{i=1}^{k} \beta_i \Delta x_i\right). \tag{6}$$

Jeżeli X_i ($i = 1, 2, \ldots, k$) jest zmienną zero-jedynkową, to e^{β_i} jest równy ilorazowi szans dla grupy, w której $X_i = 1$ oraz grupy, w której $X_i = 0$, przy pozostałych zmiennych jednakowych. Natomiast, gdy zmienna ta jest zmienną ilościową, to iloraz szans e^{β_i} mówi, jak zmieni się szansa, jeżeli zmienna X_i wzrośnie o 1 jednostkę przy pozostałych zmiennych ustalonych.

3. MODELOWANIE INTERAKCJI W REGRESJI LOGISTYCZNEJ

W literaturze istnieją różne definicje efektu interakcji. Według [6, s. 12]: „Mówimy, że istnieje efekt interakcji, kiedy wpływ zmiennej niezależnej na zmienną zależną różni się w zależności od wartości trzeciej zmiennej nazywanej zmienną moderatorem". W regresji logistycznej moderator jest więc zmienną niezależną, której wartości mają wpływ na siłę i/lub kierunek podstawowej zależności.

W modelu logitowym z dwoma predyktorami

$$\text{logit}(p) = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 \tag{7}$$

niech X_2 będzie moderatorem zmiennej X_1. Wpływ zmiennej X_1 na zmienną objaśnianą zależy więc od wartości zmiennej X_2, co oznacza, że parametr α_1 jest funkcją zmiennej X_2. Założenie, że współczynnik przy zmiennej X_1 jest liniową funkcją moderatora (np. [4] i [6])

$$\alpha_1 = \alpha_0' + \alpha_3 X_2 \tag{8}$$

prowadzi do modelu postaci

$$\text{logit}(p) = \alpha_0 + (\alpha_0' + \alpha_3 X_2) X_1 + \alpha_2 X_2 = \alpha_0 + \alpha_0' X_1 + \alpha_2 X_2 + \alpha_3 X_1 X_2, \tag{9}$$

w którym pojawia się zmienna interakcyjna będąca iloczynem zmiennych objaśniających. Zmieniając oznaczenia parametrów regresji otrzymujemy ostateczną postać modelu

$$\text{logit}(p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2. \tag{10}$$

Zgodnie z założeniem model (10) opisuje wpływ X_1 na Y, gdzie X_2 jest moderatorem, a ponieważ powyższa funkcja jest symetryczna ze względu na zmienne X_1 i X_2, więc model (10) opisuje także wpływ X_2 na Y, gdzie X_1 jest moderatorem. Modelowanie interakcji w regresji logistycznej przy użyciu iloczynu zmiennych nie wymaga identyfikacji, która zmienna niezależna jest moderatorem, a która podlega moderowaniu.

Określenie, która zmienna jest moderatorem jest arbitralne – mówi jedynie o punkcie widzenia badacza i nie ma wpływu na oszacowanie współczynników regresji, tylko na interpretację modelu.

Kontynuując powyższe rozumowanie model z trzema predyktorami uwzględniający wszystkie efekty interakcji (interakcję stopnia trzeciego oraz wszystkie interakcje stopnia drugiego) ma postać

$$\text{logit}(p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_1 X_2 + \beta_5 X_1 X_3 + \beta_6 X_2 X_3 + \beta_7 X_1 X_2 X_3. \quad (11)$$

Model zawierający wszystkie możliwe zmienne interakcyjne należy do modeli hierarchicznie dobrze zdefiniowanych (typu HWF – *hierarchically well-formulated*). Jeżeli model hierarchicznie dobrze zdefiniowany zawiera jedynie dychotomiczne zmienne niezależne, to prawdopodobieństwo wystąpienia zdarzenia oszacowane na podstawie modelu jest równe prawdopodobieństwu empirycznemu w grupach jednostek wyodrębnionych według wariantów wszystkich niezależnych zmiennych występujących w modelu. (Por. [3, s. 83])

Konstruując model należy pamiętać, że jeżeli predyktor jakościowy mający $m > 2$ wariantów jest reprezentowany przez $m - 1$ zmiennych zero-jedynkowych, to nie jest dozwolone wprowadzenie do modelu iloczynów tych zmiennych (gdyż oznaczałoby to interakcję między wariantami jednej cechy) [4].

Ocenę poprawy dopasowania modelu poprzez wprowadzenie zmiennej interakcyjnej można dokonać porównując miary dobroci dopasowania (np. miary typu pseudo-R^2) modelu zawierającego wyrażenie reprezentujące interakcję oraz modelu nie zawierającego wyrażenia reprezentującego interakcję. Weryfikację statystycznej istotności zmiennych interakcyjnych przeprowadza się w analogiczny sposób, jak pojedynczych zmiennych objaśniających [6], tzn. przy pomocy:

1. testu ilorazu wiarygodności służącego do sprawdzenia, czy wartość funkcji wiarygodności wzrosła statystycznie istotnie poprzez wprowadzenie do modelu zmiennej interakcyjnej;
2. testu Walda służącego do sprawdzenia istotności współczynnika regresji przy zmiennej interakcyjnej.

Szczególną uwagę należy zwrócić na interpretację parametru e^{β_i} dla iloczynu zmiennych. Dodatkowo obecność iloczynu zmiennych w modelu powoduje także zmianę interpretacji innych parametrów za wyjątkiem interpretacji parametru e^{β_0}, gdzie β_0 jest wyrazem wolnym. Wartość e^{β_0} jest to szansa wystąpienia zjawiska w grupie referencyjnej (wówczas wszystkie zmienne i iloczyny tych zmiennych są równe zero).

Interpretacja parametrów modelu jest odmienna dla pojedynczej zmiennej, iloczynu dwóch zmiennych, iloczynu trzech zmiennych itd., co zostanie przedstawione na przykładzie modelu (11) z trzema zmiennymi uwzględniającego wszystkie efekty interakcji. W modelu tym dla zmiennej nie będącej iloczynem np. X_1 parametr e^{β_1} można otrzymać wyznaczając iloraz szans dla jednostkowego przyrostu zmiennej X_1:

$$\psi(x_1, x_2, x_3;\ 1,\ 0,\ 0) = \frac{\gamma(x_1 + 1, x_2, x_3)}{\gamma(x_1, x_2, x_3)} = \exp(\beta_1 + \beta_4 x_2 + \beta_5 x_3 + \beta_7 x_2 x_3) \quad (12)$$

pod warunkiem, że $X_2 = X_3 = 0$ (zmienne będące moderatorami zmiennej X_1 są równe zero).

Natomiast dla iloczynu dwóch zmiennych np. X_1X_2 z modelu (11) parametr e^{β_4} można uzyskać dzieląc ilorazy szans. Jeżeli iloraz szans (12), w którym w miejsce wartości x_2 wstawiono $x_2 + 1$ zostanie podzielony przez iloraz szans (12) z drugą zmienną na poziomie x_2, to otrzymany zostanie następujący iloraz ilorazów szans:

$$\begin{aligned} &\frac{\psi(x_1, x_2 + 1,\ x_3;\ 1,\ 0,\ 0)}{\psi(x_1, x_2,\ x_3;\ 1,\ 0,\ 0)} = \\ &= \frac{\exp\left(\beta_1 + \beta_4(x_2 + 1) + \beta_5 x_3 + \beta_7(x_2 + 1)x_3\right)}{\exp\left(\beta_1 + \beta_4 x_2 + \beta_5 x_3 + \beta_7 x_2 x_3\right)} = \exp\left(\beta_4 + \beta_7 x_3\right), \end{aligned} \tag{13}$$

który przyjmuje wartość e^{β_4} dla $X_3 = 0$. Można wykazać, że $\frac{\psi(x_1, x_2 + 1,\ x_3;\ 1,\ 0,\ 0)}{\psi(x_1, x_2,\ x_3;\ 1,\ 0,\ 0)} = \frac{\psi(x_1 + 1, x_2,\ x_3;\ 0,\ 1,\ 0)}{\psi(x_1, x_2,\ x_3;\ 0,\ 1,\ 0)}$, czyli przy wyznaczaniu ilorazu ilorazów szans (13) nie ma znaczenia w jakiej kolejności badane są jednostkowe przyrosty zmiennej X_1 i X_2.

W przypadku iloczynu trzech zmiennych $X_1X_2X_3$ z modelu (11) parametr e^{β_7} można wyznaczyć jako iloraz ilorazu szans (13), w którym w miejsce wartości x_3 wstawiono $x_3 + 1$ podzielony przez iloraz ilorazu szans (13) z trzecią zmienną na poziomie x_3:

$$\frac{\dfrac{\psi(x_1, x_2 + 1,\ x_3 + 1;\ 1,\ 0,\ 0)}{\psi(x_1, x_2,\ x_3 + 1;\ 1,\ 0,\ 0)}}{\dfrac{\psi(x_1, x_2 + 1,\ x_3;\ 1,\ 0,\ 0)}{\psi(x_1, x_2,\ x_3;\ 1,\ 0,\ 0)}} = \frac{\exp\left(\beta_4 + \beta_7(x_3 + 1)\right)}{\exp\left(\beta_4 + \beta_7 x_3\right)} = \exp\left(\beta_7\right). \tag{14}$$

Podobnie, jak dla wyrażenia (13) można wykazać, że przy wyznaczaniu ilorazu ilorazów ilorazów szans (14) nie ma znaczenia w jakiej kolejności badane są jednostkowe przyrosty zmiennej X_1, X_2 i X_3.

Biorąc pod uwagę wyrażenia (12)-(14) można sformułować interpretacje parametrów e^{β} w modelach ze zmiennymi jakościowymi i ilościowymi. Interpretacja parametrów modelu w przypadku analizy interakcji między predyktorami jakościowymi reprezentowanymi przez zmienne zero-jedynkowe została przedstawiona poniżej w punktach 1)-3).

1) **Interpretacja parametru e^{β} dla zmiennej X nie będącej iloczynem zmiennych**

Parametr e^{β} jest równy ilorazowi szans dla grupy, w której $X = 1$ oraz grupy, w której $X = 0$, pod warunkiem, że zmienne będące moderatorami zmiennej X przyjmują wartość zero, czyli zmienne występujące w iloczynach z analizowaną zmienną są równe zero (przy pozostałych zmiennych jednakowych w obu grupach).

2) **Interpretacja parametru e^{β} dla zmiennej *XZ* będącej iloczynem dwóch zmiennych**

Obecność interakcji stopnia drugiego odzwierciedla się w zróżnicowaniu ilorazów szans. Porównania ilorazów szans można dokonać wyznaczając ich iloraz i sprawdzając, czy różni się statystycznie istotnie od 1. Jeżeli zmienna Z zostanie potraktowana jako moderator zmiennej X, to współczynnik e^{β} jest równy ilorazowi dwóch ilorazów szans otrzymanemu poprzez podzielenie ilorazu szans dla grupy, w której $X = 1$ oraz grupy, w której $X = 0$ pod warunkiem, że $Z = 1$ przez iloraz tych szans pod warunkiem, że $Z = 0$ (przy pozostałych moderatorach zmiennej X równych zero oraz ustalonej wartości pozostałych zmiennych niezależnych).

3) **Interpretacja parametru e^{β} dla zmiennej *QXZ* będącej iloczynem trzech zmiennych**

Tak, jak obecność interakcji stopnia drugiego można stwierdzić dzieląc ilorazy szans, tak obecność interakcji stopnia trzeciego można zaobserwować dzieląc ilorazy ilorazów szans. Niech zmienne Q oraz Z będą moderatorami zmiennej X, to wartość e^{β} można uzyskać dzieląc iloraz ilorazu szans otrzymany tak, jak przy badaniu interakcji stopnia drugiego między zmienną X oraz Z, lecz pod warunkiem, że $Q = 1$ przez iloraz ilorazu szans otrzymany analogicznie, tym razem pod warunkiem, że $Q = 0$ (przy pozostałych moderatorach zmiennej X równych zero oraz ustalonej wartości pozostałych zmiennych niezależnych).

W sytuacji, gdy analizowana jest interakcja między predyktorami jakościowymi i ilościowymi lub interakcja między predyktorami tylko ilościowymi interpretacje 1)-3) należy nieco zmienić. W przypadku zmiennej ilościowej w powyższych interpretacjach zastępuje się grupy zakodowane jako 0 oraz jako 1 przez dwie grupy, dla których zmienna ilościowa różni się o 1 jednostkę. (por. (12)-(14) i [6]).

4. ANALIZA RYZYKA ZGONU Z WYKORZYSTANIEM MODELU LOGITOWEGO

Analiza dotyczy ryzyka zgonu osób starszych (w wieku od 60 lat) w województwie pomorskim w roku 2009. Model zbudowano na podstawie następujących danych pochodzących z GUS:

1. liczba osób zmarłych w roku 2009,
2. liczba osób żyjących na końcu roku 2008 oraz 2009

sklasyfikowanych jednocześnie według roku urodzenia, wieku ukończonego, płci, miejsca zamieszkania (miasto/wieś). Wiek osób przyjęto na moment 31 XII 2008, czyli w roku 2009 obserwowane były generacje osób urodzonych w roku 1948 i wcześniej, przy czym uwzględniono poprawkę na liczbę osób migrujących (por. [12, s. 36-37]). W okresie roku kalendarzowego obserwacji poddano 379 tys. mieszkańców województwa pomorskiego. Jednoczesna klasyfikacja osób według wszystkich cech badania po-

zwoliła na uzyskanie indywidualnych danych dla każdej jednostki.[5] Parametry modelu wyznaczono metodą największej wiarygodności przy użyciu modułu „Uogólnione modele liniowe" zawartego w pakiecie komputerowym *Statistica 8.0*.

Zróżnicowanie ryzyka zgonu w zależności od płci, wieku, miejsca zamieszkania można wstępnie zaobserwować porównując empiryczne (nie poddane modelowaniu) wartości prawdopodobieństwa oraz szansy zgonu w grupach osób sklasyfikowanych pod względem tych cech (patrz tabela 1).

Tabela 1

Empiryczne wartości prawdopodobieństwa oraz szansy zgonu według płci, miejsca zamieszkania i 5-letnich grup wieku w województwie pomorskim w 2009 roku

Wiek	Prawdopodobieństwo zgonu				Szansa zgonu			
	Miasta		Wieś		Miasta		Wieś	
	Mężczyźni	Kobiety	Mężczyźni	Kobiety	Mężczyźni	Kobiety	Mężczyźni	Kobiety
60-64	0,01957	0,00890	0,02257	0,01075	0,01996	0,00898	0,02309	0,01087
65-69	0,03073	0,01393	0,03274	0,01605	0,03170	0,01413	0,03385	0,01632
70-74	0,03997	0,02106	0,05221	0,02248	0,04164	0,02151	0,05508	0,02300
75-79	0,06355	0,03681	0,07222	0,04058	0,06786	0,03821	0,07784	0,04230
80-84	0,09416	0,06926	0,10776	0,07484	0,10394	0,07442	0,12078	0,08089
85+	0,16845	0,13944	0,18120	0,15395	0,20257	0,16203	0,22130	0,18197

Źródło: obliczenia własne na podstawie danych GUS.

Regresja logistyczna umożliwiła znalezienie statystycznie istotnych predyktorów ryzyka zgonu oraz odkrycie efektów interakcji między tymi predyktorami. Zmienna objaśniana przyjęła postać

$$Y = \begin{cases} 1 & \text{zgon} \\ 0 & \text{przezycie} \end{cases} . \tag{15}$$

Dwa predyktory jakościowe płeć i miejsce zamieszkania zostały wprowadzone do modelu przy pomocy zmiennych zero-jedynkowych. Grupę referencyjną utworzyły kobiety mieszkające w mieście. Zmienną wiek w analizie uwzględniono na dwa sposoby:

[5] Przy użyciu metody największej wiarygodności model estymowany na podstawie danych indywidualnych jest identyczny z modelem estymowanym na podstawie danych pogrupowanych jednocześnie według wszystkich wariantów cech występujących dla danych indywidualnych [7, s. 105-106].

1. jako zmienną skategoryzowaną – wyodrębniono sześć kategorii wieku (patrz tabela 1)[6] reprezentowanych w modelu przez pięć zmiennych zero-jedynkowych[7] (najmłodsza grupa wieku od 60 do 64 lat została grupą referencyjną),
2. jako zmienną ciągłą minus 60 (w tym przypadku 60-latkowie utworzyli grupę referencyjną).

Powyższe dwa podejścia mają swoje ograniczenia. Wprowadzenie predyktora wiek do modelu jako zmiennej ilościowej wymaga założenia, aby logit był linową funkcją wieku, a tym samym oznacza to, że każdy wzrost wieku o rok powoduje taki sam procentowy wzrost szansy zgonu (stały iloraz szans dla jednakowych przyrostów wieku). Założenie to można ominąć kategoryzując zmienną i wprowadzając ją do modelu jako zespół zmiennych zero-jedynkowych. Dla każdej zmiennej zero-jedynkowej oszacowany zostaje odrębny współczynnik regresji, więc ilorazy szans w sąsiednich grupach wieku nie muszą być stałe. Jednakże wyodrębnienie grup wieku powoduje utratę szczegółowości danych.

4.1 WYNIKI ESTYMACJI MODELU LOGITOWEGO BEZ UWZGLĘDNIENIA EFEKTÓW INTERAKCJI

Konstrukcja dwumianowego modelu logitowego pozwoliła na identyfikację predyktorów ryzyka zgonu. Wszystkie analizowane zmienne objaśniające okazały się statystycznie istotne zarówno w wariancie z wiekiem jako zmienną skategoryzowaną (tabela 2) oraz w wariancie z wiekiem jako zmienną ciągłą (tabela 3). Występowanie zróżnicowania ryzyka zgonu w zależności od płci, wieku, miejsca zamieszkania można stwierdzić sprawdzając, czy współczynnik regresji β różni się statystycznie istotnie od 0 na poziomie istotności $\alpha = 0,05$, co jest równoznaczne z tym, że iloraz szans równy w modelu logitowym e^{β} różni się statystycznie istotnie od 1, a także tym, że 95% przedział ufności dla e^{β} nie zawiera 1 (por. np. [5]).

W przypadku wyrazu wolnego współczynnik e^{β} równa się szansie wystąpienia zgonu w grupie referencyjnej, tzn. grupie kobiet w wieku 60-64 lata (tabela 2) lub w wieku 60 lat (tabela 3) zamieszkujących w miastach (por. wartości z tabeli 2 i 3 z rzeczywistą wartością z tabeli 1). Jest to grupa najniższego ryzyka w badanej populacji (potwierdzają to dodatnie współczynniki regresji β przy zmiennych objaśniających). Analizując ilorazy szans e^{β} z tablicy 2 i 3 można stwierdzić, że oba modele wskazują, że szansa zgonu mężczyzn jest większa o około 71%-73% niż szansa zgonu kobiet, natomiast szansa zgonu mieszkańca wsi jest większa o około 14% niż szansa zgonu mieszkańca miasta. Dla zmiennej skategoryzowanej wiek oszacowano ilorazy szans zgonu e^{β} w danej grupie wieku w stosunku do grupy referencyjnej 60-64 lata, np.

[6] Są to grupy wieku stosowane tradycyjnie do prezentacji danych demograficznych. Problem wyodrębnienia grup wieku, które są zróżnicowane ze względu na prawdopodobieństwo zgonu jest złożonym tematem (dyskusji można poddać wybór liczby grup, czy granic przedziałów wieku, tzw. punktów odcięcia).

[7] Inny sposób kodowania grup wieku zastosowano m. in. w [9, s. 148-151].

Tabela 2

Wyniki estymacji modelu logitowego z zero-jedynkowymi zmiennymi objaśniającymi bez uwzględnienia efektów interakcji

	Ocena parametru β	Błąd standardowy	e^{β}	95% przedział dla e^{β}	
Wyraz wolny	-4,5421	0,0281	0,0107	0,0101	0,0113
Płeć (1-mężczyzna, 0-kobieta)	0,5381	0,0174	1,7127	1,6552	1,7722
Miejsce (1-wieś, 0-miasto)	0,1324	0,0191	1,1415	1,0997	1,1850
Wiek 65-69	0,4341	0,0354	1,5435	1,4401	1,6545
Wiek 70-74	0,7888	0,0334	2,2007	2,0612	2,3496
Wiek 75-79	1,2922	0,0321	3,6409	3,4190	3,8771
Wiek 80-84	1,8565	0,0322	6,4014	6,0098	6,8185
Wiek 85+	2,6186	0,0320	13,7166	12,8833	14,6037
test ilorazu wiarygodności: *p-value* = 0,0000 pseudo-R^2 Vealla-Zimmermanna = 0,1040			czułość modelu = 58,9% swoistość modelu = 74,7%		

Źródło: obliczenia własne z wykorzystaniem programu *Statistica*.

Tabela 3

Wyniki estymacji modelu logitowego z zero-jedynkowymi zmiennymi płeć i miejsce zamieszkania oraz zmienną ciągłą wiek bez uwzględnienia efektów interakcji

	Ocena parametru β	Błąd standardowy	e^{β}	95% przedział dla e^{β}	
Wyraz wolny	-4,8650	0,0227	0,0077	0,0074	0,0081
Płeć (1-mężczyzna, 0-kobieta)	0,5467	0,0175	1,7275	1,6694	1,7877
Miejsce (1-wieś, 0-miasto)	0,1358	0,0191	1,1454	1,1033	1,1891
Wiek	0,1001	0,0010	1,1053	1,1031	1,1075
test ilorazu wiarygodności: *p-value* = 0,0000 pseudo-R^2 Vealla-Zimmermanna = 0,1070			czułość modelu = 66,3% swoistość modelu = 67,8%		

Źródło: obliczenia własne z wykorzystaniem programu *Statistica*.

szansa zgonu w grupie 65-69 lat jest większa o 54% niż w grupie 60-64 lata, szansa zgonu w grupie 70-74 lat jest większa o 120% niż w grupie 60-64 lata itd. Na tej podstawie można także obliczyć, iż w stosunku do sąsiedniej młodszej grupy szansa zgonu jest większa w grupie: 70-74 lat o 43%, 75-79 lat o 65%, 80-84 lat o 76%, 85+ lat o 114%. Natomiast w modelu przedstawionym w tabeli 3 zwiększenie wieku o rok powoduje wzrost szansy zgonu o 10,5%, czyli zwiększenie wieku o 5 lat powoduje wzrost szansy zgonu o 65% ($e^{5\beta} = 1,6496$).

4.2 WYNIKI ESTYMACJI MODELU Z UWZGLĘDNIENIEM EFEKTÓW INTERAKCJI

Płeć, miejsce zamieszkania i wiek są istotnymi predyktorami ryzyka zgonu, jednakże ich wpływ na zmienną objaśnianą może zależeć od wzajemnych interakcji. Uwzględnienie wszystkich interakcji (do stopnia 3 włącznie) w modelu ze wszystkimi zmiennymi zero-jedynkowymi powoduje, że prawdopodobieństwa zgonu oszacowane na podstawie takiego modelu są równe prawdopodobieństwom empirycznym dla danych pogrupowanych jednocześnie według płci, miejsca zamieszkania i grup wieku (wartości zmiennych objaśniających podzieliły w tym przypadku zbiorowość na $2 \times 2 \times 6 = 24$ grupy). Prowadzi to do rozbudowanego modelu (z 24 parametrami) mimo występowania tylko trzech predyktorów. Jednak nie wszystkie zmienne interakcyjne okazały się statystycznie istotne. W tabeli 4 przedstawiono model ze zmiennymi dychotomicznymi zawierający istotne statystycznie efekty interakcji. Budowa modelu metodą krokową

Tabela 4
Wyniki estymacji modelu logitowego z zero-jedynkowymi zmiennymi objaśniającymi z uwzględnieniem efektów interakcji

	Ocena parametru β	Błąd standardowy	e^{β}	95% przedział dla e^{β}	
Wyraz wolny	-4,6966	0,0437	0,0091	0,0084	0,0099
Płeć (1-mężczyzna, 0-kobieta)	0,7877	0,0538	2,1984	1,9783	2,4430
Miejsce (1-wieś, 0-miasto)	0,1287	0,0190	1,1374	1,0957	1,1806
Wiek 65-69	0,4411	0,0591	1,5545	1,3845	1,7453
Wiek 70-74	0,8407	0,0545	2,3181	2,0830	2,5796
Wiek 75-79	1,4243	0,0514	4,1550	3,7565	4,5957
Wiek 80-84	2,0859	0,0498	8,0519	7,3027	8,8780
Wiek 85+	2,8733	0,0486	17,6951	16,0863	19,4648
Płeć * Wiek 65-69	-0,0017	0,0738	0,9983	0,8638	1,1538
Płeć * Wiek 70-74	-0,0637	0,0691	0,9383	0,8194	1,0744
Płeć * Wiek 75-79	-0,2033	0,0662	0,8161	0,7168	0,9290
Płeć * Wiek 80-84	-0,4352	0,0668	0,6472	0,5677	0,7377
Płeć * Wiek 85+	-0,5717	0,0676	0,5646	0,4945	0,6446
test ilorazu wiarygodności: *p-value* = 0,0000 pseudo-R^2 Vealla-Zimmermanna = 0,1055			czułość modelu = 64,9% swoistość modelu = 68,9%		

Źródło: obliczenia własne z wykorzystaniem programu *Statistica*.

postępującą oraz krokową wsteczną dała ten sam rezultat.[8] Statystycznie istotna okazała się interakcja płci z wiekiem, natomiast nieistotne okazały się interakcje stopnia drugiego miejsca zamieszkania z płcią oraz miejsca zamieszkania z wiekiem, jak też interakcja stopnia trzeciego między analizowanymi trzema predyktorami.[9]

Współczynnik e^{β} dla zmiennej płeć (tabela 4) mówi, że szansa zgonu mężczyzn jest około dwa razy większa niż szansa zgonu kobiet, ale tylko w wieku 60-64 lata (grupa referencyjna dla moderatora zmiennej płeć) zarówno na wsi, jak i w mieście (miejsce zamieszkania nie jest moderatorem zmiennej płeć). Interpretacja współczynnika e^{β} dla zmiennej miejsce zamieszkania nie zmieniła się, gdyż zmienna ta nie posiada moderatora (szansa zgonu mieszkańca wsi jest większa o około 14% niż szansa zgonu mieszkańca miasta). W przypadku zmiennych zero-jedynkowych reprezentujących grupy wieku współczynnik e^{β} jest równy ilorazowi szans zgonu w danej grupie wieku oraz grupie referencyjnej w wieku 60-64 lata, ale tylko dla kobiet (moderator zmiennej wiek równa się wówczas zero), np. szansa zgonu kobiet w wieku 65-69 lat jest o 55% większa niż kobiet w wieku 60-64 lata (zarówno na wsi, jak i w mieście). Współczynnik e^{β} przy iloczynie zmiennej płeć oraz zmiennej zero-jedynkowej reprezentującej grupę wieku interpretowany jest jako iloraz dwóch ilorazów szans, np. iloraz szans zgonu mężczyzn i kobiet w wieku 85 lat i więcej jest mniejszy o 43,5% od ilorazu szans mężczyzn i kobiet w wieku 60-64 lata (iloraz szans zgonu mężczyzn i kobiet zmniejsza się z wiekiem, gdyż powyżej 60 lat maleje przewaga umieralności mężczyzn nad umieralnością kobiet).

W celu zobrazowania dopasowania modeli ze wszystkimi zmiennymi zero-jedynkowymi na wykresie 1 przedstawiono różnice względne (w %) między prawdopodobieństwem zgonu teoretycznym a prawdopodobieństwem empirycznym we wszystkich grupach wieku dla kobiet i mężczyzn mieszkających w miastach.[10] Model bez interakcji jest znacznie gorzej dopasowany: w zależności od wariantów analizowanych cech różnice względne wahały się od kilku do około 20%. W przypadku modelu ze wszystkimi interakcjami stopnia drugiego oraz modelu z interakcjami wyłącznie statystycznie istotnymi (tabela 4) różnice względne są znacznie mniejsze i nie przekraczają 5%, więc przyjęcie modelu zawierającego interakcje jedynie statystycznie istotne jest całkowicie zadawalające.[11]

---

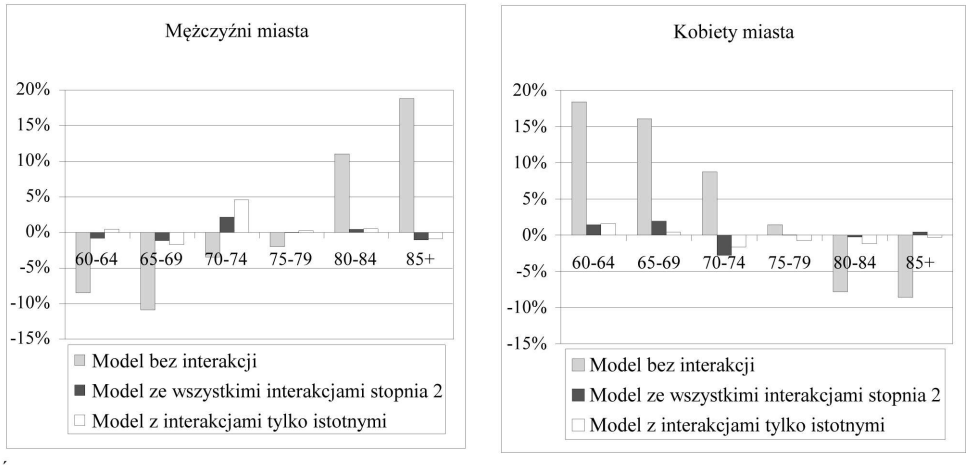
[8] Brak reguł przy przeszukiwaniu zbioru zmiennych może prowadzić do przypadkowych wyników, dlatego wykorzystuje się algorytmy do wyboru zmiennych objaśniających do modelu (*stepwise logistic regression*). (Por. [2] i [5])

[9] Zgodnie z kryterium informacyjnym Akaike'a oraz bayesowskim kryterium informacyjnym spośród modeli ze zmiennymi dychotomicznymi najlepszy okazał się również model opisany w tabeli 4.

[10] Dla osób mieszkających na wsi prawidłowości zaobserwowane w odchyleniach wartości teoretycznych od empirycznych były podobne jak dla mieszkańców miast (nieistotne interakcje miejsca zamieszkania z pozostałymi predyktorami).

[11] Na wykresie 1 nie zaprezentowano odchyleń dla modelu HWF ze wszystkimi efektami interakcji (w tym przypadku do stopnia trzeciego włącznie), gdyż dla takiego modelu odchylenia są równe zero (patrz uwaga w punkcie 3).

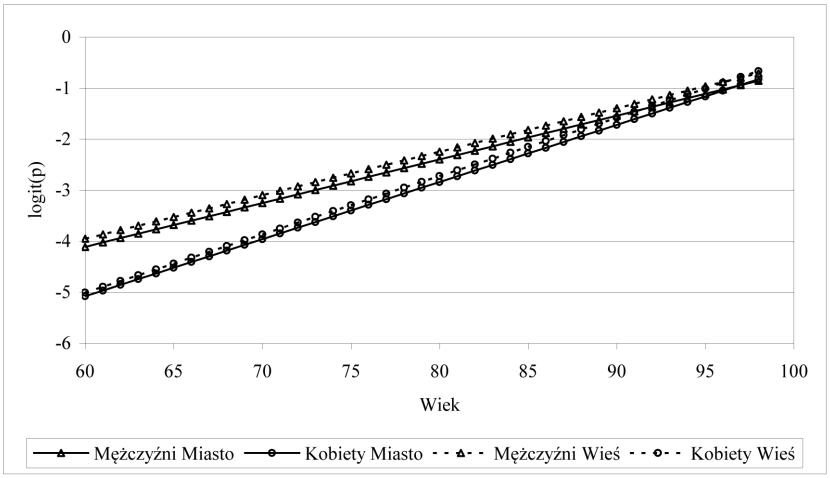
Wykres 1. Różnice względne w % między prawdopodobieństwem zgonu teoretycznym (oszacowanym na podstawie modelu z zero-jedynkowymi zmiennymi objaśniającymi) a prawdopodobieństwem empirycznym



Źródło: opracowanie własne.

W przypadku wprowadzenia do modelu predyktora wiek jako zmiennej ciągłej można graficznie przedstawić logit jako liniową funkcję wieku w grupach według płci i miejsca zamieszkania. Wykres 2 prezentuje przebieg tej funkcji oszacowanej na podstawie modelu zawierającego wszystkie efekty interakcji (czyli do stopnia trzeciego włącznie).

Wykres 2. Logit jako funkcja zmiennej ciągłej wiek w grupach według płci i miejsca zamieszkania – wyniki estymacji modelu zawierającego wszystkie efekty interakcji



Źródło: opracowanie własne.

Różne kąty nachylenia prostych w grupie kobiet i mężczyzn świadczą o występowaniu interakcji między predyktorami wiek i płeć. Natomiast zbliżony do równoległego przebieg prostych dla mężczyzn zamieszkałych w miastach i na wsi oraz równoległy przebieg prostych dla kobiet zamieszkałych w miastach i na wsi świadczy o nieistotnej sile interakcji między predyktorami wiek i miejsce zamieszkania (por. [4] i [6]). Zbliżone odległości między parami tych prostych świadczą także o niewielkiej sile interakcji między predyktorami płeć i miejsce zamieszkania. Procedura konstrukcji modelu zarówno krokowa postępująca oraz krokowa wsteczna potwierdziły powyższe obserwacje. Statystycznie istotne okazały się predyktory płeć, miejsce zamieszkania, wiek oraz iloczyn wieku i płci (tabela 5).[12]

Tabela 5

Wyniki estymacji modelu logitowego z zero-jedynkowymi zmiennymi płeć i miejsce zamieszkania oraz zmienną ciągłą wiek z uwzględnieniem efektów interakcji

	Ocena parametru β	Błąd standardowy	e^{β}	95% przedział dla e^{β}	
Wyraz wolny	-5,0962	0,0297	0,0061	0,0058	0,0065
Płeć (1-mężczyzna, 0-kobieta)	0,9927	0,0383	2,6984	2,5030	2,9090
Miejsce (1-wieś, 0-miasto)	0,1332	0,0191	1,1425	1,1005	1,1860
Wiek	0,1126	0,0014	1,1191	1,1160	1,1222
Wiek*Płeć	-0,0270	0,0021	0,9734	0,9695	0,9773
test ilorazu wiarygodności: *p-value* = 0,0000 pseudo-R^2 Vealla-Zimmermanna = 0,1088			czułość modelu = 68,1% swoistość modelu = 66,1%		

Źródło: obliczenia własne z wykorzystaniem programu *Statistica*.

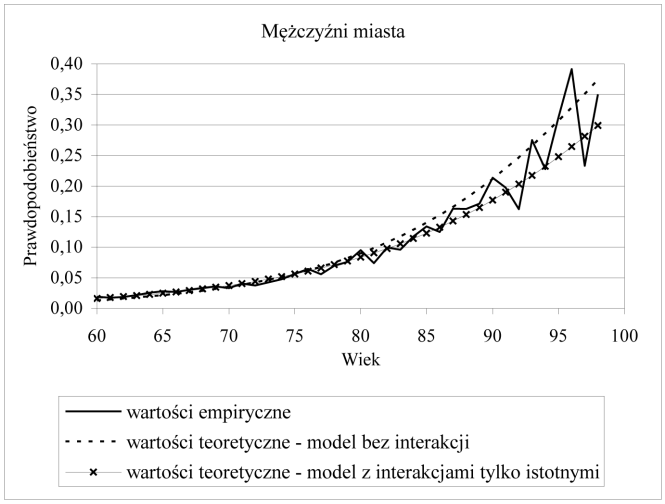
Współczynniki e^{β} dla zmiennej płeć (tabela 5) mówią, że szansa zgonu mężczyzn jest około 2,7 razy większa niż szansa zgonu kobiet, ale tylko w wieku 60 lat (grupa referencyjna dla moderatora zmiennej płeć) zarówno na wsi, jak i w mieście (miejsce zamieszkania nie jest moderatorem zmiennej płeć). Interpretacja współczynnika e^{β} dla zmiennej miejsce zamieszkania nie zmieniła się, gdyż zmienna ta nie posiada moderatora (szansa zgonu mieszkańca wsi jest większa o około 14% niż szansa zgonu mieszkańca miasta). W przypadku zmiennej wiek współczynnik e^{β} informuje, że wraz ze wzrostem wieku o jeden rok szansa zgonu rośnie o około 12%, ale tylko w grupie kobiet (moderator zmiennej wiek równa się wówczas zero). Współczynnik e^{β} przy iloczynie zmiennej płeć oraz wiek mówi, że jeżeli wiek wzrośnie o jeden rok, to iloraz szans zgonu mężczyzn i kobiet zmaleje o 2,7% (wraz z wiekiem maleje przewaga umieralności mężczyzn nad umieralnością kobiet).

Prawdopodobieństwa zgonu empiryczne i teoretyczne jako funkcje wieku dla mężczyzn mieszkających w miastach przedstawiono na wykresie 3, a dla kobiet mieszkają-

[12] Kryterium informacyjne Akaike'a oraz bayesowskie kryterium informacyjne również wskazały na model opisany w tabeli 5.

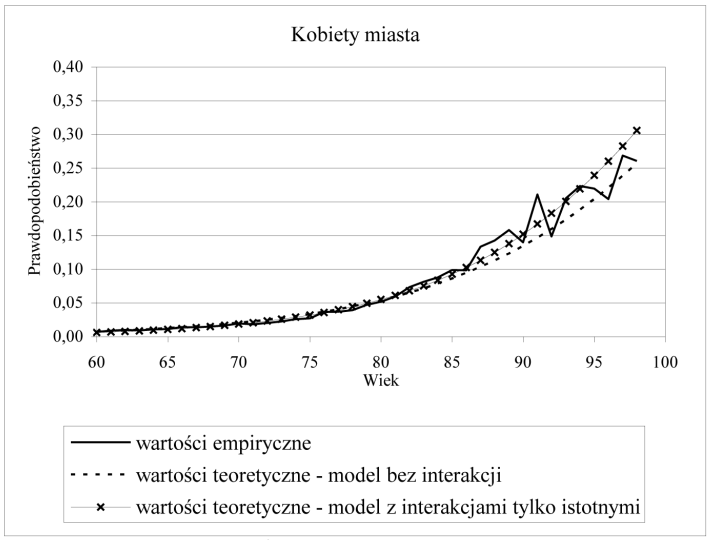
cych w miastach na wykresie 4.[13] Na wykresie dokonano porównania dla modelu bez interakcji oraz modelu z interakcjami tylko statystycznie istotnymi.

Wykres 3. Prawdopodobieństwa zgonu empiryczne i teoretyczne (oszacowane na podstawie modelu ze zmienną ciągłą wiek) w przypadku mężczyzn zamieszkujących w miastach



Źródło: opracowanie własne.

Wykres 4. Prawdopodobieństwa zgonu empiryczne i teoretyczne (oszacowane na podstawie modelu ze zmienną ciągłą wiek) w przypadku kobiet zamieszkujących w miastach



Źródło: opracowanie własne.

[13] Analogiczne prawidłowości można zaobserwować dla mieszkańców wsi (nieistotne interakcje miejsca zamieszkania z pozostałymi predyktorami).

Empiryczne prawdopodobieństwa zgonu traktowane jako funkcja wieku charakteryzują się nieregularnym przebiegiem, w szczególności dla najstarszych roczników (mniejsza liczba danych). Do ich wygładzenia (wyrównania) można wykorzystać modele logitowe. W przypadku modelu bez interakcji znaki odchyleń wartości empirycznych od teoretycznych tworzą zbyt długie serie. W przypadku mężczyzn model bez interakcji przeszacowuje prawdopodobieństwa zgonu (długie serie punktów empirycznych pod wykresem oszacowanej funkcji (wykres 3)), natomiast w przypadku kobiet model bez interakcji niedoszacowuje prawdopodobieństwa zgonu (długie serie punktów empirycznych nad wykresem oszacowanej funkcji (wykres 4))[14]. Uwzględnienie interakcji powoduje, że odchylenia wartości empirycznych od teoretycznych można uznać za losowe[15].

Ponieważ część efektów interakcji okazała się statystycznie nieistotna, więc przebieg funkcji ze wszystkimi zmiennymi interakcyjnymi oraz z wybranymi zmiennymi interakcyjnymi jest zbliżony. W praktyce poszukuje się funkcji jak najlepiej dopasowanej i jednocześnie o możliwie najmniejszej liczbie parametrów, więc jako model analityczny (parametryczny) ludzkiego procesu wymieralności wystarczy przyjąć model jedynie z interakcjami statystycznie istotnymi.

5. PODSUMOWANIE

W niniejszym artykule przedstawiono propozycję zastosowania modelu logitowego w analizie zróżnicowania ryzyka zgonu. Regresja logistyczna pozwoliła na oszacowanie prawdopodobieństwa zgonu w zależności od poziomu zmiennych objaśniających: płeć, miejsce zamieszkania oraz wiek. Zaletą modelu logitowego jest to, że obok ilościowych zmiennych niezależnych, mogą występować w modelu także jakościowe zmienne niezależne zakodowane w odpowiedni sposób.

W zależności od przeznaczenia modelu zmienną ilościową można skategoryzować, a następnie otrzymane kategorie zakodować np. za pomocą zmiennych zerojedynkowych. Model z wiekiem jako zmienną skategoryzowaną (tabela 2 i 4) pozwala na oszacowanie odrębnych współczynników regresji dla grup wieku, co pozwala uniknąć założenia o stałym ilorazie szans dla jednakowych przyrostów wieku i może być cenne dla oceny ryzyka np. metodą scoringową. Natomiast model z wiekiem jako zmienną ciągłą (tabela 3 i 5) może znaleźć zastosowanie do wygładzania ciągu empirycznych prawdopodobieństw zgonu (patrz wykres 3 i 4).

Wprowadzenie zmiennych interakcyjnych w postaci iloczynu zmiennych pozwoliło na wykrycie efektów interakcji między zmiennymi objaśniającymi, przy czym statystycznie istotna okazała się interakcja płci z wiekiem, natomiast nieistotne okazały

[14] W modelu bez interakcji uwzględniony jest stały (uśredniony) poziom nadumieralności mężczyzn w stosunku do kobiet, lecz nadumieralność mężczyzn maleje po 60 roku życia.

[15] Innym sposobem rozwiązania tego problemu byłoby oszacowanie modeli odrębnych dla mężczyzn i kobiet. Jednakże budowa modelu logitowego ze zmiennymi interakcyjnymi pozwala dodatkowo na ocenę siły i kierunku efektu interakcji.

się interakcje stopnia drugiego miejsca zamieszkania z płcią oraz miejsca zamieszkania z wiekiem, jak też interakcja stopnia trzeciego między analizowanymi trzema predyktorami.

Szczególną uwagę w artykule zwrócono na możliwość interpretacji parametrów e^{β}. Interpretacja ta zależy od sposobu kodowania zmiennych oraz od tego, czy współczynnik regresji β występuje przy zmiennej nie będącej iloczynem, czy też przy zmiennej będącej iloczynem, jak też od tego ile zmiennych występuje w iloczynie. Do interpretacji parametru e^{β} oszacowanego dla zmiennej nie będącej iloczynem wykorzystuje się pojęcie ilorazu szans, dla zmiennej będącej iloczynem dwóch zmiennych wykorzystuje się pojęcie ilorazu ilorazu szans, dla zmiennej będącej iloczynem trzech zmiennych wykorzystuje się pojęcie ilorazu ilorazu ilorazu szans itd. W artykule rozważone zostały przypadki interakcji między predyktorami jakościowymi, między predyktorem jakościowym i ilościowym oraz między predyktorami ilościowymi.

Model logitowy pozwala na uwzględnienie cech demograficzno-społecznych oraz efektu interakcji między tymi cechami przy szacowaniu prawdopodobieństwa zgonu, co może znaleźć praktyczne zastosowania, np. w dziedzinie ubezpieczeń, planach emerytalnych, czy medycynie.

Uniwersytet Gdański

LITERATURA

[1] Agresti A., [2002], *Categorical Data Analysis*, John Wiley & Sons, New Jersey.
[2] Christensen R. Ch., [1997], *Log-linear models and logistic regression*, Springer, New York.
[3] Gruszczyński M., [2010], *Modele zmiennych jakościowych dwumianowych*, W: *Mikroekonometria. Modele i metody analizy danych indywidualnych,* pod red. Gruszczyński M., Wolters Kluwer Polska, Warszawa.
[4] Harrell F., [2001], *Regression Modeling Strategies with Applications to Linear Models, Logistic Regression, and Survival Analysis*, Springer-Verlag, New York.
[5] Hosmer D., Lemeshow S., [2000], *Applied Logistic Regression*, John Wiley & Sons, New Jersey.
[6] Jaccard J., [2001], *Interaction Effects in Logistic Regression*, Sage University Papers, Series: „Quantitative Applications in the Social Sciences", 07-135, Thousand Oaks.
[7] Jong de P., Heller G. Z., [2008], *Generalized Linear Models for Insurance Data*, Cambridge University Press, Cambridge.
[8] Maddala G. S., [1983], *Limited-dependent and qualitative variables in econometrics*, Cambridge University Press, Cambridge.
[9] Mazurek E., [2000], *Model logitowy*, W: *Metody oceny i porządkowania ryzyka w ubezpieczeniach życiowych*, pod red. Ostasiewicz S., Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.
[10] McCullagh P., Nelder J. A., [1989], *Generalized Linear Models*, Chapman & Hall, London.
[11] Tabeau E., Willekens F., van Poppel F., [2002], *Parameterization as a tool in analyzing age, period and cohort effects on mortality: A case study of the Netherlands*, W: *The Life Table. Modelling Survival and Death*, pod red. Wunsch G., Mouchart M., Duchene J., Kluwer Academic Publishers, Dordrecht.
[12] *Trwanie życia w 2009 r.*, [2010], Główny Urząd Statystyczny, Warszawa.

EFEKTY INTERAKCJI MIĘDZY ZMIENNYMI OBJAŚNIAJĄCYMI W MODELU LOGITOWYM W ANALIZIE ZRÓŻNICOWANIA RYZYKA ZGONU

S t r e s z c z e n i e

Celem pracy jest identyfikacja predyktorów ryzyka zgonu oraz zbadanie efektów interakcji pomiędzy nimi. W artykule wykorzystano model regresji logistycznej do oszacowania prawdopodobieństw zgonu osób starszych (w wieku od 60 lat) w województwie pomorskim w 2009 roku. Jako predyktory ryzyka zgonu przyjęto: wiek, płeć oraz miejsce zamieszkania (miasto/wieś). W modelu uwzględniono wiek na dwa sposoby: jako zmienną ciągłą oraz jako zmienną skategoryzowaną. Przeprowadzono analizę efektów interakcji między predyktorami poprzez wprowadzenie do modelu iloczynu zmiennych objaśniających. Szczególną uwagę zwrócono na interpretację współczynników w modelu zawierającym zmienne interakcyjne. Rozważono przypadki interakcji między predyktorami jakościowymi, między predyktorem jakościowym i ilościowym oraz między predyktorami ilościowymi. Statystycznie istotna okazała się interakcja płci z wiekiem.

Słowa kluczowe: regresja logistyczna, efekt interakcji, prawdopodobieństwo zgonu

INTERACTION EFFECTS BETWEEN PREDICTOR VARIABLES IN A LOGISTIC MODEL IN AN ANALYSIS OF THE DIVERSITY OF DEATH RISK

S u m m a r y

The aims of this paper include the identification of predictors of death risk and the examination of interaction effects between them. In this study, a logistic regression model is used to estimate death probability at old age (above 60) in the Pomorskie Voivodship in 2009. The following risk factors of death are considered: age, gender and place of residence (urban/rural areas). In the model, age is treated both as a continuous variable and as a categorical variable. The paper presents an analysis of interaction effects between predictors with the use of product terms in the logistic regression model. The emphasis is on the interpretation of the coefficients of the interactive logistic model. The study includes cases of interactions between qualitative predictors, between qualitative and quantitative predictors, and between quantitative predictors. It appears that the interaction between gender and age is statistically significant.

Key words: logistic regression, interaction effect, death probability

PRZEGLĄD STATYSTYCZNY
R. LVIII – ZESZYT 1-2 – 2011

JOANNA OLBRYŚ

OBCIĄŻENIE ESTYMATORA WSPÓŁCZYNNIKA ALFA JENSENA A INTERPRETACJE PARAMETRÓW KLASYCZNYCH MODELI MARKET-TIMING[1]

1. WSTĘP

W 1972 roku E. Fama [10] zaproponował dychotomiczny podział umiejętności zarządzającego portfelem, czyli rozróżnienie umiejętności w zakresie prognozowania:

- w skali mikro (w literaturze określane mianem selektywności aktywów),
- w skali makro (w literaturze określane jako market-timing, czyli tzw. wyczucie rynku).

W swojej pracy Fama powołał się m.in. na artykuł Michaela C. Jensena *The performance of mutual funds in the period 1945 – 1964* [19], który ukazał się w 1968 r. w czasopiśmie „The Journal of Finance". W artykule tym Jensen wprowadził współczynnik α, zwany w literaturze przedmiotu alfą Jensena, jako miarę umiejętności zarządzającego portfelem inwestycyjnym w zakresie selektywności aktywów. Jensen przedstawił w skrócie przekształcenia, na podstawie których wywnioskował, że estymator współczynnika alfa jest obciążony dodatnio, co oczywiście ma istotny wpływ na interpretację wartości tego parametru. Wnioski Jensena odnośnie interpretacji można podsumować następującym stwierdzeniem: estymator $\hat{\alpha}$ może być dodatni z dwóch powodów: 1) dodatkowych zysków osiąganych przez menadżera wykorzystującego swoje umiejętności w zakresie doboru aktywów; 2) dodatniego obciążenia estymatora parametru α.

W roku 1977, również w „The Journal of Finance′′,ukazał się z kolei artykuł Dwighta Granta *Portfolio performance and the „cost" of timing decisions* [14], w którym autor podważa prawdziwość wzoru przedstawionego przez Jensena pisząc wprost o błędzie matematycznym w przekształceniach [14, str. 837]. Prezentuje własne wyprowadzenie wzoru na wartość oczekiwaną odpowiedniego estymatora i podaje nową interpretację, według której estymator współczynnika alfa jest obciążony ujemnie.

Wprowadza to sporo zamieszania w literaturze przedmiotu, ponieważ jedni autorzy powołują się w swoich pracach na interpretację Jensena (np. [5], [11], [13], [15], [17], [28]), natomiast inni na interpretację Granta (np. [6], [14]). Niektórzy autorzy komentują obie interpretacje (np. [31]). Niestety, w konsekwencji wnioski dotyczą-

[1] Praca naukowa finansowana ze środków na naukę w latach 2009 – 2011 jako projekt badawczy własny Nr N N113 173237 „Modele market-timing wspomagające ocenę efektywności zarządzania portfelami polskich funduszy inwestycyjnych"

ce wyników badań empirycznych na rynkach funduszy inwestycyjnych, w zakresie umiejętności stosowania przez zarządzających funduszami technik market-timing oraz selektywności aktywów, nie są jednoznaczne. Głównym punktem artykułu jest wyprowadzenie spornego wzoru, w celu wyciągnięcia ostatecznych wniosków dotyczących statystycznego oddziaływania obciążenia estymatora współczynnika alfa Jensena na pozostałe parametry modelu regresji. Dodatkowo przedstawione zostaną przykłady klasycznych modeli market-timing polskich funduszy inwestycyjnych akcji, jak również zostanie zbadana stabilność parametrów otrzymanych modeli ekonometrycznych w okresie styczeń 2003 – grudzień 2010, w celu potwierdzenia wyników estymacji oraz interpretacji.

2. WSPÓŁCZYNNIK ALFA – WZÓR JENSENA (1968)

Głównym celem jednej z najważniejszych prac Jensena [19] było wprowadzenie miary efektów zarządzania portfelem (*portfolio performance measure).*

Zgodnie z modelem CAPM, wartość oczekiwana stopy zwrotu z papieru wartościowego może być oszacowana na podstawie wzoru [19, str.390]:

$$E\left(R_j\right) = R_F + \beta_j \cdot [E\left(R_M\right) - R_F] \tag{1}$$

gdzie:

$E\left(R_j\right)$ – wartość oczekiwana stopy zwrotu z akcji j-tej,

$E\left(R_M\right)$ – wartość oczekiwana stopy zwrotu z portfela rynkowego M,

R_F – stopa zwrotu wolna od ryzyka,

$\beta_j = \dfrac{cov\left(R_j, R_M\right)}{\sigma^2\left(R_M\right)}$ – współczynnik beta Sharpe'a akcji j-tej.

Jensen pisze, powołując się na artykuł E. Famy z 1968r. [9], że wartość współczynnika β_j w równaniu (1) jest w przybliżeniu równa wartości współczynnika b_j w „modelu rynku" postaci [19, str. 391]:

$$R_{j,t} = E\left(R_{j,t}\right) + b_j \cdot \pi_t + e_{j,t} \tag{2}$$

gdzie π_t jest czynnikiem rynku (nieobserwowalnym).

Zakładamy, że zmienne losowe π_t oraz $e_{j,t}$ są niezależne, o rozkładzie normalnym, czyli spełnione są warunki: $E\left(\pi_t\right) = 0$; $E\left(e_{j,t}\right) = 0$;

$cov\left(\pi_t, e_{j,t}\right) = 0$; $cov\left(e_{j,t}, e_{i,t}\right) = \begin{cases} 0, j \neq i \\ \sigma^2\left(e_j\right), j = i \end{cases}$.

Współczynniki β_j oraz b_j we wzorach (1) i (2) nie posiadają indeksu t, czyli przyjmujemy ich niezmienność w czasie.

Aproksymacja stopy zwrotu z portfela rynkowego ma postać [19, str. 392]:

$$R_{M,t} \cong E\left(R_{M,t}\right) + \pi_t \tag{3}$$

zatem: $E\left(R_{M,t}\right) = R_{M,t} - \pi_t$.

Z (1) oraz (3) otrzymujemy:

$$E\left(R_{j,t}\right) + \beta_j \cdot \pi_t + e_{j,t} \cong R_{F,t} + \beta_j \cdot \left(R_{M,t} - \pi_t - R_{F,t}\right) + \beta_j \cdot \pi_t + e_{j,t} \tag{4}$$

Po podstawieniu (2) oraz redukcji uzyskamy:

$$R_{j,t} - R_{F,t} = \beta_j \cdot \left(R_{M,t} - R_{F,t}\right) + e_{j,t} \tag{5a}$$

$$r_{j,t} = \beta_j \cdot r_{M,t} + e_{j,t} \tag{5b}$$

gdzie:

$r_{j,t} = R_{j,t} - R_{F,t}$ – nadwyżka stopy zwrotu z akcji j-tej nad wolną od ryzyka stopą zwrotu w okresie t,

$r_{M,t} = R_{M,t} - R_{F,t}$ – nadwyżka stopy zwrotu z portfela rynkowego M nad wolną od ryzyka stopą zwrotu w okresie t.

Równania (5a)-(5b) nie odzwierciedlają sytuacji, w której zarządzający portfelem jest „doskonałym prognostą" (*superior forecaster*) i w celu uzyskania jak najwyższej stopy zwrotu będzie wybierał papiery wartościowe o $e_{j,t} > 0$. Aby uwzględnić w równaniach (5a)-(5b) taką możliwość można wprowadzić stałą α_j różną od zera:

$$R_{j,t} - R_{F,t} = \alpha_j + \beta_j \cdot \left(R_{M,t} - R_{F,t}\right) + u_{j,t} \tag{6a}$$

$$r_{j,t} = \alpha_j + \beta_j \cdot r_{M,t} + u_{j,t} \tag{6b}$$

gdzie: $u_{j,t}$ - składnik losowy o $E\left(u_{j,t}\right) = 0$.

Działania prowadzone przez zarządzającego portfelem wymagają wykorzystania umiejętności w zakresie:

1. przewidywania zmian cen pojedynczych aktywów, czyli selektywności papierów wartościowych,
2. przewidywania zmian na rynku, globalnie (czyli zmian czynnika rynkowego π_t w modelu rynku (2)).

Jeśli menadżer posiada umiejętności w zakresie selektywności aktywów, to $\alpha_j > 0$.

W celu uwzględnienia zmienności w czasie parametru beta, Jensen proponuje przedstawienie $\beta_{j,t}$ jako zmiennej losowej [19, str. 395]:

$$\beta_{j,t} = \beta_j + \varepsilon_{j,t} \tag{7}$$

gdzie:

β_j – docelowy średni poziom ryzyka (*target risk level*),

$\varepsilon_{j,t}$ – składnik losowy o $E\left(\varepsilon_{j,t}\right) = 0$.

Jeśli inwestor potrafi nawet w pewnym stopniu przewidzieć kierunek zmian na rynku, to fakt ten można zapisać formalnie w postaci zależności pomiędzy składnikiem losowym $\varepsilon_{j,t}$ z równania (7) oraz zmienną π_t (równanie (2)):

$$\varepsilon_{j,t} = a_j \cdot \pi_t + \omega_{j,t} \tag{8}$$

gdzie: $\omega_{j,t}$ – składnik losowy o rozkładzie normalnym oraz $E\left(\omega_{j,t}\right) = 0$.

Współczynnik a_j w równaniu (8) jest dodatni – jeśli menadżer ma zdolność przewidywania w skali makro; jest równy zero – jeśli menadżer nie ma takich zdolności (lub ich nie wykorzystuje). Przypadku $a_j < 0$ nie rozważamy.

Po podstawieniu zależności z równania (8) do równania (7) otrzymamy:

$$\beta_{j,t} = \beta_j + \varepsilon_{j,t} = \beta_j + a_j \cdot \pi_t + \omega_{j,t} \tag{9}$$

Równanie (9) przedstawia model regresji – zależność między zmienną $\beta_{j,t}$ a czynnikiem rynkowym π_t, z uwzględnieniem składnika losowego $\omega_{j,t}$.

Podstawiając z kolei zależność (7) do równań (6a)-(6b) uzyskamy:

$$R_{j,t} - R_{F,t} = \alpha_j + \beta_{j,t} \cdot (R_{M,t} - R_{F,t}) + u_{j,t} \tag{10a}$$

$$r_{j,t} = \alpha_j + \beta_{j,t} \cdot r_{M,t} + u_{j,t} \tag{10b}$$

Jeśli estymator parametru beta (estymator średniego poziomu ryzyka) jest nieobciążony, to nieobciążony jest również estymator parametru alfa (estymator miary umiejętności w zakresie selektywności aktywów). Jensen pisze, że można pokazać, iż przy założeniu, że składnik losowy $\omega_{j,t}$ nie jest skorelowany ze zmienną π_t, wartość oczekiwana estymatora MNK parametru beta wynosi [19, str. 396]:

$$E\left(\hat{\beta}_j\right) = \frac{cov\left[\left(R_{j,t} - R_{F,t}\right), (R_{M,t} - R_{F,t})\right]}{\sigma^2 (R_M)} = \beta_j - a_j \cdot E\,(R_M) \tag{11}$$

Na podstawie wzoru (11) Jensen wnioskuje, że:

1. jeśli inwestor nie ma umiejętności przewidywania w skali mikro, czyli $a_j = 0$, wtedy $E\left(\hat{\beta}_j\right) = \beta_j$ - estymator parametru beta jest nieobciążony,
2. jeśli inwestor ma umiejętności przewidywania w zakresie selektywności aktywów, czyli $a_j > 0$, wtedy $E\left(\hat{\beta}_j\right) \neq \beta_j$, czyli obserwujemy obciążenie estymatora parametru beta.

Zatem, jeśli estymator parametru beta jest obciążony ujemnie, to estymator parametru alfa jest obciążony dodatnio, aby spełniony był warunek, że prosta regresji (10b) przechodzi przez punkt o współrzędnych $(\overline{r_M}, \overline{r_j})$.

3. WYPROWADZENIE WZORU

Praca Jensena [19] nie zawiera dowodu zależności (11), jedynie wzmiankę, jak można ten wzór uzyskać. W celu uproszczenia zapisu (a co za tym idzie, również dowodu), wzór (11) można przedstawić w innej postaci, z wykorzystaniem oznaczeń dotyczących nadwyżek stóp zwrotu (*excess returns)* $r_{j,t} = R_{j,t} - R_{F,t}$ oraz $r_{M,t} = R_{M,t} - R_{F,t}$ odpowiednio, opuszczając jednocześnie indeks t:

$$E\left(\hat{\beta}_j\right) = \frac{cov\left[\left(R_j - R_F\right), \left(R_M - R_F\right)\right]}{\sigma^2\left(r_M\right)} = \frac{cov\left(r_j, r_M\right)}{\sigma^2\left(r_M\right)} \tag{12}$$

Wprowadzając oznaczenia:

$r_j = \beta_j \cdot r_M$ – ze wzoru (5b),

$a_j = \dfrac{cov\left(\beta_j, r_M\right)}{\sigma^2\left(r_M\right)}$ – na podstawie równania (9), przyjmując za czynnik rynkowy nadwyżkę stopy zwrotu z portfela rynkowego M,

oraz korzystając z własności wartości oczekiwanej, kowariancji i wariancji zmiennych losowych (w szczególności o rozkładzie normalnym dwumianowym) otrzymujemy:

$$\begin{aligned}
E\left(\hat{\beta}_j\right) &= \frac{cov\left[\left(R_j - R_F\right), \left(R_M - R_F\right)\right]}{\sigma^2\left(r_M\right)} = \frac{cov\left(r_j, r_M\right)}{\sigma^2\left(r_M\right)} = \frac{cov\left(\beta_j \cdot r_M, r_M\right)}{\sigma^2\left(r_M\right)} = \\
&= \frac{E\left(\beta_j \cdot r_M^2\right) - E\left(\beta_j \cdot r_M\right) \cdot E\left(r_M\right)}{\sigma^2\left(r_M\right)} = \\
&= \frac{cov\left(\beta_j, r_M^2\right) + E\left(\beta_j\right) \cdot E\left(r_M^2\right) - \left[cov\left(\beta_j, r_M\right) + E\left(\beta_j\right) \cdot E\left(r_M\right)\right] \cdot E\left(r_M\right)}{\sigma^2\left(r_M\right)} = \\
&= \frac{cov\left(\beta_j, r_M^2\right) + E\left(\beta_j\right) \cdot \left[\sigma^2\left(r_M\right) + E^2\left(r_M\right)\right] - cov\left(\beta_j, r_M\right) \cdot E\left(r_M\right) - E\left(\beta_j\right) \cdot E^2\left(r_M\right)}{\sigma^2\left(r_M\right)} = \\
&= \frac{cov\left(\beta_j, r_M^2\right) + E\left(\beta_j\right) \cdot \sigma^2\left(r_M\right) - cov\left(\beta_j, r_M\right) \cdot E\left(r_M\right)}{\sigma^2\left(r_M\right)} = \\
&= E\left(\beta_j\right) + \frac{cov\left(\beta_j, r_M^2\right) - cov\left(\beta_j, r_M\right) \cdot E\left(r_M\right)}{\sigma^2\left(r_M\right)} = E\left(\beta_j\right) - a_j \cdot E\left(r_M\right)
\end{aligned}$$

Uzyskana w wyniku przekształceń postać jest zgodna z odpowiednim wzorem z pracy Jensena [19, str. 396].

4. WSPÓŁCZYNNIK ALFA – WNIOSKI GRANTA (1977)

D. Grant w pracy [14] podważył wnioski Jensena dotyczące interpretacji parametru alfa twierdząc, że w przypadku portfeli inwestycyjnych zarządzanych z wykorzystaniem

technik market-timing i selektywności aktywów, wartość oczekiwana estymatora parametru beta jest zawyżona, a nie zaniżona (jak sugerował Jensen), a co za tym idzie, estymator parametru alfa jest obciążony ujemnie (a nie dodatnio, jak u Jensena). Grant stwierdził wprost, że w oryginalnej pracy Jensena jest błąd matematyczny [14, str. 837].

Grant wprowadził następujące pomocnicze oznaczenia, które wykorzystał w działaniach:

$$\Delta r_M = r_M - E(r_M); \quad \Delta\beta = \beta - E(\beta); \quad r = \beta \cdot r_M; \quad a = \frac{cov(\beta, r_M)}{\sigma^2(r_M)}.$$

W wyniku przekształceń Grant otrzymał [14, str. 843]:

$$E\left(\hat{\beta}\right) = \frac{cov(r, r_M)}{\sigma^2(r_M)} = E(\beta) + \frac{cov(\beta, r_M) \cdot E(r_M)}{\sigma^2(r_M)} = E(\beta) + a \cdot E(r_M) \qquad (13)$$

Wzory uzyskane przez Jensena oraz Granta różnią się znakiem wyrażenia $a \cdot E(r_M)$:

- $E\left(\hat{\beta}\right) = E(\beta) - a \cdot E(r_M)$ – wzór Jensena (11);
- $E\left(\hat{\beta}\right) = E(\beta) + a \cdot E(r_M)$ – wzór Granta (13).

Jak zostało udowodnione w rozdziale 3, wzór Jensena jest prawidłowy, natomiast wzór Granta – błędny.

5. WSPÓŁCZYNNIK ALFA JENSENA W KLASYCZNYCH MODELACH MARKET-TIMING – ZASTOSOWANIE I INTERPRETACJA

W ostatnich latach w literaturze przedmiotu pojawiło się wiele różnorodnych modeli tzw. wyczucia rynku, czyli market-timing.

Modele te w większości wywodzą się od modeli podstawowych, będąc ich modyfikacjami (np. [2], [6], [11], [12], [13], [28], [31]).

Jeden z pierwszych i najprostszych modeli parametrycznych do badania umiejętności w zakresie stosowania technik market-timing oraz selektywności aktywów przez zarządzających portfelami inwestycyjnymi opracowali Treynor i Mazuy w 1966r. [38]. Zdefiniowali oni umiejętność wyczucia rynku jako właściwą reakcję na zmiany stopy zwrotu z portfela rynkowego. Pożądane jest, aby ryzyko portfela było wyższe w okresie występowania dodatniej rynkowej stopy zwrotu w porównaniu z okresami spadkowymi. W celu testowania umiejętności zarządzającego portfelem w zakresie wyczucia rynku można zastosować MNK do estymacji parametrów strukturalnych następującego modelu regresji [38] (oznaczanego dalej jako model T-M):

$$R_{P,t} = \alpha_P + \beta_P \cdot R_{M,t} + \gamma_P \cdot R_{M,t}^2 + \varepsilon_{P,t} \qquad (14)$$

gdzie:

$R_{P,t}$ – stopa zwrotu z portfela P w okresie t,

$R_{M,t}$ – stopa zwrotu z portfela rynkowego M w okresie t,

α_P – miara umiejętności zarządzającego portfelem P w zakresie selektywności aktywów (czyli alfa Jensena),

β_P – miara ryzyka systematycznego portfela P,

γ_P – miara umiejętności stosowania techniki market-timing przez zarządzającego portfelem.

Model T-M można szacować również z wykorzystaniem nadwyżek stóp zwrotu portfela P oraz portfela rynkowego M, w stosunku do stopy zwrotu wolnej od ryzyka. Przyjmie wtedy postać:

$$r_{P,t} = \alpha_P + \beta_P \cdot r_{M,t} + \gamma_P \cdot r_{M,t}^2 + \varepsilon_{P,t} \tag{15}$$

$r_{P,t} = R_{P,t} - R_{F,t}$ – nadwyżka stopy zwrotu z portfela P nad wolną od ryzyka stopą zwrotu w okresie t,

$r_{M,t} = R_{M,t} - R_{F,t}$– nadwyżka stopy zwrotu z portfela rynkowego M nad wolną od ryzyka stopą zwrotu w okresie t,

Pozostałe oznaczenia jak we wzorze (14).

Interpretacja wartości oszacowania współczynnika α_P w równaniu (15) jest następująca: jeśli menadżer ma zdolność przewidywania w zakresie selektywności aktywów, to współczynnik α_P jest dodatni oraz istotny statystycznie. Zgodnie z interpretacją Jensena, dodatnia, ale nieistotna statystycznie wartość estymatora tego parametru może być efektem dodatniego obciążenia estymatora i niekoniecznie musi świadczyć o umiejętnościach zarządzającego portfelem.

Jeśli menadżer zarządzający portfelem np. funduszu inwestycyjnego zwiększa (zmniejsza) ekspozycję portfela na ryzyko rynkowe w przypadku wzrostów (spadków) stopy zwrotu z portfela rynkowego, wtedy stopa zwrotu z portfela jest wypukłą funkcją rynkowej stopy zwrotu i parametr γ_P jest dodatni. Wielkość tego parametru świadczy o stopniu umiejętności stosowania techniki market-timing.

Przykład 1

Badanie umiejętności zarządzających portfelami funduszy inwestycyjnych w zakresie selektywności aktywów oraz tzw. wyczucia rynku przeprowadzono z wykorzystaniem próby statystycznej zawierającej dzienne obserwacje stopy zwrotu 15 wybranych otwartych funduszy inwestycyjnych akcji z okresu od stycznia 2003 do grudnia 2010 (po 2013 pomiarów dla każdego z funduszy). Analogicznie, dzienne dane głównego indeksu giełdy warszawskiej WIG wykorzystano jako stopy zwrotu z portfela rynkowego. Z kolei średnia dzienna stopa rentowności bonów skarbowych 52 – tygodniowych została użyta jako wolna od ryzyka stopa zwrotu. Wyniki estymacji modeli T-M (15) wybranych funduszy, z wykorzystaniem estymatorów odpornych Neweya-Westa (*HAC*), przedstawia tabela 1. W przypadku 7 z 15 funduszy uzyskano dodatnie i istotne statystycznie oszacowania parametru α_P Jensena. Zgodnie z interpretacją Jensena, może to świadczyć o posiadaniu przez zarządzających portfelami funduszy umiejętności

w zakresie selektywności aktywów, natomiast może być również efektem dodatniego obciążenia estymatora tego parametru.

Dodatkowo można zinterpretować uzyskane wartości oszacowań pozostałych parametrów modeli. Wartość estymatora parametru β_P, reprezentującego w modelach T-M ryzyko systematyczne portfeli funduszy, wahała się w przedziale [0,41; 0,81] i była we wszystkich przypadkach statystycznie istotna. Średnia wartość tego parametru wyniosła 0,64. Z kolei wartość estymatora parametru γ_P, będącego miarą umiejętności stosowania techniki market-timing przez zarządzającego portfelem, była istotnie mniejsza od zera w przypadku 12 z 15 funduszy, co w literaturze przedmiotu jest interpretowane jako niepokojąca informacja o braku takich umiejętności lub nie wykorzystywaniu ich przez zarządzających portfelami funduszy (np. [2], [6], [12], [17], [20], [31]). Średnia wartość estymatora parametru γ_P w całej grupie funduszy była równa w badanym okresie – 1,39.

Tabela 1

Model T-M (15), dane dzienne z okresu styczeń 2003 – grudzień 2010

	Fundusz akcji	$\hat{\alpha}_P$	$\hat{\beta}_P$	$\hat{\gamma}_P$	$\overline{R^2}$
1	Arka BZ WBK FIO Subfundusz Arka Akcji	0,0006***	0,72***	-2,10***	0,63
2	Aviva Investors FIO Subfundusz Aviva Polskich Akcji	0,0005***	0,76***	-1,67***	0,70
3	BPH FIO Parasolowy Subfundusz BPH Akcji	0,0001	0,73***	-0,79	0,73
4	DWS Polska FIO Top 25 Małych Spółek	0,0005*	0,41***	-2,34***	0,26
5	DWS Polska FIO Akcji Dużych Spółek	0,0002	0,66***	-1,15	0,41
6	DWS Polska FIO Akcji Plus	0,0003	0,56***	-1,44*	0,38
7	ING Parasol FIO Subfundusz Akcji	0,0001	0,76***	-0,74*	0,72
8	Legg Mason Akcji FIO	0,0003***	0,70***	-0,92*	0,72
9	Millennium FIO Subfundusz Akcji	0,0001	0,69***	-1,03*	0,69
10	Novo FIO Subfundusz Novo Akcji	0,0004**	0,51***	-1,95*	0,32
11	Pioneer FIO Subfundusz Pioneer Akcji Polskich	0,0001	0,81***	-1,33*	0,71
12	PKO Akcji - FIO	0,0003	0,57***	-2,11*	0,45
13	PZU FIO Akcji KRAKOWIAK	0,0002	0,71***	-1,24**	0,71
14	Skarbiec FIO Subfundusz Akcji Skarbiec - Akcja	0,0003*	0,48***	-0,61	0,31
15	UniFundusze FIO Subfundusz UniKorona Akcje	0,0005**	0,52***	-1,36*	0,31
	średnia	**0,0003**	**0,64**	**-1,39**	**0,54**

Źródło: opracowanie własne (z wykorzystaniem pakietu *Gretl 1.8.5*)

Parametry istotnie różniące się od zera:
* istotność na poziomie 0,1; ** istotność na poziomie 0,05; *** istotność na poziomie 0,01.

Kolejny model tzw. wyczucia rynku, czyli test parametryczny Henrikssona-Mertona z 1984r. ([16], [17]) umożliwia identyfikację i ocenę umiejętności zarządzającego portfelem inwestycyjnym w zakresie stosowania techniki market-timing, jak również selekcji aktywów. Model H-M ma postać [16, str. 527]:

$$r_{P,t} = \alpha_P + \beta_{1P} \cdot x_t + \beta_{2P} \cdot y_t + \varepsilon_{P,t} \tag{16}$$

gdzie:
$x_t = r_{M,t}$ – nadwyżka stopy zwrotu z portfela rynkowego nad stopą zwrotu wolną od ryzyka,
$y_t = \max\{0, -x_t\}$,
α_P – miara umiejętności zarządzającego portfelem P w zakresie selektywności aktywów (czyli alfa Jensena),
β_{1P} – miara ryzyka systematycznego portfela P,
β_{2P} – miara umiejętności stosowania techniki market-timing przez zarządzającego portfelem,
$\varepsilon_{P,t}$ – składnik losowy, spełniający założenia modelu CAPM: $E(\varepsilon_{P,t}) = 0$, $E(\varepsilon_{P,t}|x_t) = 0$, $E\left(\varepsilon_{P,t}\left|\varepsilon_{P,t-i}\right.\right) = 0; i = 1, 2, 3, ...$

Równanie (16) wynika z modelu Mertona [22]. Merton uzasadnił, że doskonałe wyczucie rynku można teoretycznie uzyskać budując strategię opcyjną typu „*protective put*", inwestując część gotówki w portfel rynkowy, za pozostałą zaś część nabywając „darmowe" opcje sprzedaży (z ceną realizacji $R_{F,t}$) każdej złotówki znajdującej się w portfelu rynkowym. Strategia ta replikuje strukturę stóp zwrotu uzyskaną w wyniku perfekcyjnego stosowania strategii market-timing. Zgodnie z modelem Mertona, estymator $\hat{\alpha}_P$ jest miarą wpływu umiejętności menadżera w zakresie doboru papierów wartościowych na wyniki inwestycyjne. Testowana hipoteza zerowa ma postać:

$$H_0 : \alpha_P = 0 \tag{17}$$

tzn. przypuszczamy, że zarządzający portfelem nie posiada umiejętności w zakresie selektywności aktywów, czyli przewidywania w skali mikro.

Estymator $\hat{\beta}_{1P}$ reprezentuje część środków zainwestowaną w portfel rynkowy zgodnie ze strategią opcyjną Mertona, zaś estymator $\hat{\beta}_{2P}$ – liczbę "darmowych" opcji sprzedaży. W tym kontekście, badanie umiejętności w zakresie wyczucia rynku jest równoznaczne z testowaniem hipotezy zerowej:

$$H_0 : \beta_{2P} = 0 \tag{18}$$

czyli zarządzający portfelem nie posiada umiejętności w zakresie wyczucia rynku lub ich nie wykorzystuje. Ujemna wartość estymatora $\hat{\beta}_{2P}$ oznacza negatywny wpływ techniki market-timing na wartość portfela.

Przykład 2

Analogicznie, jak w przykładzie 1, badanie umiejętności zarządzających portfelami funduszy inwestycyjnych w zakresie selektywności aktywów oraz wyczucia rynku przeprowadzono z wykorzystaniem próby statystycznej zawierającej dzienne obserwacje stopy zwrotu 15 wybranych otwartych funduszy inwestycyjnych akcji z okresu od stycznia 2003 do grudnia 2010. Wyniki estymacji modeli H-M (16) wybranych funduszy, z wykorzystaniem estymatorów odpornych Neweya-Westa (*HAC*) prezentuje tabela 2. W przypadku 12 funduszy odrzucono wprawdzie hipotezę zerową (17), dotyczącą parametru α_P Jensena, ale istotnie różne od zera wartości tego parametru były i tak bardzo małe. Wnioski mogą być zatem analogiczne, jak w przykładzie 1. Zgodnie z interpretacją Jensena, dodatnia wartość estymatora parametru α_P może świadczyć o posiadaniu przez zarządzających portfelami funduszy umiejętności w zakresie selektywności aktywów, ale może też wynikać z dodatniego obciążenia tego estymatora.

Tabela 2

Model H-M (16), dane dzienne z okresu styczeń 2003 – grudzień 2010

	Fundusz akcji	$\hat{\alpha}_P$	$\hat{\beta}_P$	$\hat{\gamma}_P$	$\overline{R^2}$
1	Arka BZ WBK FIO Subfundusz Arka Akcji	0,001***	0,64***	-0,17***	0,63
2	Aviva Investors FIO Subfundusz Aviva Polskich Akcji	0,0009***	0,69***	-0,14***	0,70
3	BPH FIO Parasolowy Subfundusz BPH Akcji	0,0004*	0,69***	-0,08*	0,73
4	DWS Polska FIO Top 25 Małych Spółek	0,001***	0,31***	-0,20***	0,26
5	DWS Polska FIO Akcji Dużych Spółek	0,0004	0,61***	-0,09	0,41
6	DWS Polska FIO Akcji Plus	0,0006*	0,50***	-0,13*	0,38
7	ING Parasol FIO Subfundusz Akcji	0,0003	0,73***	-0,08*	0,72
8	Legg Mason Akcji FIO	0,0006***	0,66***	-0,09**	0,72
9	Millennium FIO Subfundusz Akcji	0,0004**	0,64***	-0,11**	0,69
10	Novo FIO Subfundusz Novo Akcji	0,0009**	0,43***	-0,16**	0,32
11	Pioneer FIO Subfundusz Pioneer Akcji Polskich	0,0004	0,76***	-0,12**	0,71
12	PKO Akcji – FIO	0,0008**	0,48***	-0,18**	0,45
13	PZU FIO Akcji KRAKOWIAK	0,0005**	0,66***	-0,11***	0,71
14	Skarbiec FIO Subfundusz Akcji Skarbiec – Akcja	0,0006*	0,44***	-0,07	0,31
15	UniFundusze FIO Subfundusz UniKorona Akcje	0,0009**	0,45***	-0,13*	0,31
	średnia	**0,001**	**0,58**	**-0,12**	**0,54**

Źródło: opracowanie własne (z wykorzystaniem pakietu *Gretl 1.8.5*)

Parametry istotnie różniące się od zera:

* istotność na poziomie 0,1; ** istotność na poziomie 0,05; *** istotność na poziomie 0,01.

Interpretacje wartości oszacowań pozostałych parametrów otrzymanych modeli są następujące. Wartość estymatora parametru β_{1P}, reprezentującego w modelach H-M ryzyko systematyczne portfeli funduszy, należała do przedziału [0,31; 0,76] i była we wszystkich przypadkach istotna statystycznie. Przeciętna wartość tego parametru wyniosła 0,58 i była nieznacznie niższa w porównaniu z wynikiem w tabeli 1. Z kolei wartość estymatora parametru β_{2P}, będącego miarą umiejętności stosowania techniki market-timing przez zarządzającego portfelem, była istotnie mniejsza od zera w przypadku 13 spośród 15 funduszy. Odrzucenie hipotezy zerowej (18) i ujemne wartości tego estymatora mogą świadczyć o braku umiejętności w zakresie stosowania techniki market-timing lub nie wykorzystywaniu ich przez zarządzających portfelami funduszy. Wartość średnia estymatora parametru β_{2P} wyniosła w badanym okresie -0,12.

6. TESTOWANIE STABILNOŚCI PARAMETRÓW MODELI MARKET-TIMING

Jednym z testów stabilności parametrów modelu ekonometrycznego, zbudowanego w oparciu o szereg czasowy, jest test *CUSUM (CUmulated SUM of Residuals),* oparty o tzw. reszty rekursywne [21]. Autorami testu są Brown, Durbin i Evans [3]. Test ten można stosować nawet wtedy, gdy nie znamy momentu zmiany strukturalnej (czyli tzw. punktu zwrotnego) procesu i nie zakładamy, że ona wystąpi. Zatem jest on bardziej uniwersalny, niż np. predykcyjny test Chowa. Załóżmy, że estymowany model regresji ma postać [3, str. 150-151]:

$$y_t = x_t^T \cdot \beta_t + \varepsilon_t, \;\; t = 1, \ldots, T \tag{19}$$

gdzie:

y_t jest obserwacją zmiennej zależnej w okresie t,

x_t jest kolumnowym wektorem obserwacji k zmiennych niezależnych w okresie t; pierwsza zmienna niezależna x_{1t} jest równa 1 dla każdego t, jeśli model zawiera wyraz wolny,

β_t jest kolumnowym wektorem parametrów w okresie t,

ε_t jest składnikiem losowym w okresie t;

przyjmujemy, że składniki losowe mają rozkład normalny ze średnią zero i wariancją σ_t^2, $t=1,\ldots,T$.

Testowana hipoteza zerowa ma postać [3, str. 151]:

$$\begin{aligned} H_0 : \beta_1 = \beta_2 = \cdots = \beta_T = \beta \\ \sigma_1^2 = \sigma_2^2 = \cdots \sigma_T^2 = \sigma^2 \end{aligned} \tag{20}$$

Reszta rekursywna o numerze t jest błędem predykcji wartości zmiennej zależnej y_t, gdy estymacja modelu odbywa się z wykorzystaniem $(t-1)$ obserwacji:

$$e_t = y_t - x_t^T \cdot b_{t-1} \tag{21}$$

gdzie:

b_{t-1} jest wektorem współczynników (regresja na podstawie $(t-1)$ obserwacji).

Wariancja reszty rekursywnej jest równa:

$$\sigma_{pt}^2 = \sigma^2 \cdot [1 + x_t^T \cdot (X_{t-1}^T \cdot X_{t-1})^{-1} \cdot x_t] \tag{22}$$

gdzie:

$X_{t-1}^T = [x_1, x_2, \ldots, x_{t-1}]$, t=k+1,...,T.

Przeskalowaną resztę rekursywną, czyli stosunek reszty rekursywnej e_t do obciążenia prognozy w okresie t oznaczamy przez w_t i obliczamy ze wzoru:

$$w_t = \frac{e_t}{\sqrt{1 + x_t^T \cdot (X_{t-1}^T \cdot X_{t-1})^{-1} \cdot x_t}} \tag{23}$$

Test *CUSUM* tworzony jest na podstawie prób od $(k+1)$ do T. Test ten oparty jest na wykresie skumulowanych sum przeskalowanych reszt rekursywnych postaci:

$$W_t = \frac{1}{s} \cdot \sum_{j=k+1}^{j=t} w_j \tag{24}$$

gdzie:

$s^2 = \frac{1}{T-k-1} \cdot \sum_{j=k+1}^{T} (w_j - \overline{w})^2$ jest wariancją przeskalowanych reszt rekursywnych,

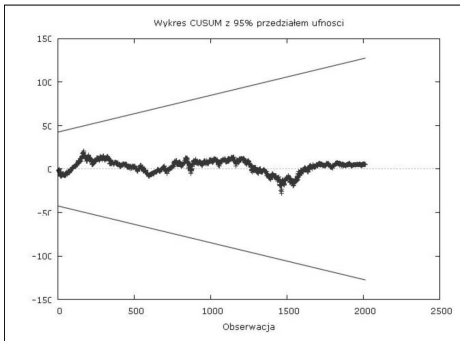
$s = \sqrt{s^2}$;

$\overline{w} = \frac{1}{T-k} \cdot \sum_{j=k+1}^{T} w_j$ jest średnią arytmetyczną przeskalowanych reszt rekursywnych, obliczonych ze wzoru (23).

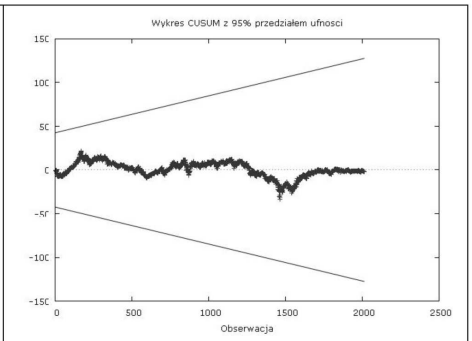
Idea testu polega na wyznaczeniu pary prostych leżących symetrycznie poniżej i ponad prostą $E(W_t) = 0$ tak, aby prawdopodobieństwo przekroczenia jednej lub obu linii wynosiło α, gdzie α jest wymaganym poziomem istotności. Jeśli suma reszt rekursywnych, określona wzorem (24), przekracza na wykresie górną lub dolną linię krytyczną, to można wnioskować o wystąpieniu punktu zwrotnego (inaczej: zmiany strukturalnej) w danym momencie. Oznacza to odrzucenie hipotezy zerowej (20), czyli model nie jest stabilny w badanym okresie. Linie krytyczne są to dwie proste łączące punkty o współrzędnych: $(k;\ \pm a \cdot \sqrt{T-k})$ i $(T;\ \pm 3a \cdot \sqrt{T-k})$, odpowiednio, gdzie a jest parametrem, którego wartość uzależniona jest od ustalonego poziomu istotności α. Najczęściej używane w praktyce pary wartości a oraz α to [3, str. 153-154]: (α=0,01 a=1,143), (α=0,05, a=0,948) oraz (α=0,10, a=0,850).

W celu potwierdzenia wyników estymacji, weryfikacji oraz interpretacji dokonano testowania stabilności parametrów otrzymanych klasycznych modeli T-M (15) oraz H-M (16) w okresie 2.01.2003r.–31.12.2010r., z wykorzystaniem uniwersalnego testu *CUSUM,* z 95% przedziałem ufności (α=0,05, a=0,948). Wykresy testu wybranych 6

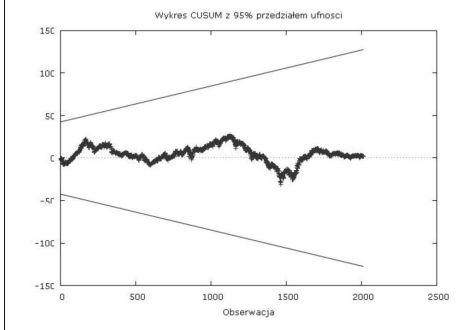
spośród 15 modeli market-timing badanych funduszy akcji polskich przedstawia Rysunek 1. Zaprezentowano trzy „najlepsze" (fundusze: DWS, Novo, Unikorona) oraz trzy „najgorsze" (fundusze: DWS25, ING, Pioneer) wykresy. Ze względu na ograniczoną objętość niniejszej pracy, wykresy testu *CUSUM* pozostałych modeli są dostępne na życzenie. W przypadku żadnego z funduszy nie odrzucono hipotezy zerowej (20), co potwierdza stałość w czasie parametrów strukturalnych wszystkich modeli poddanych analizie.



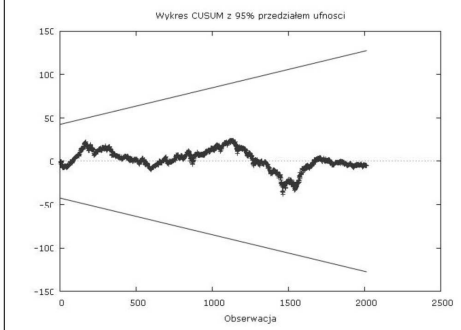
Fundusz DWS model T-M (15)



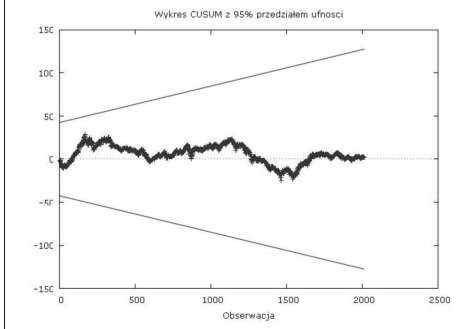
Fundusz DWS model H-M (16)



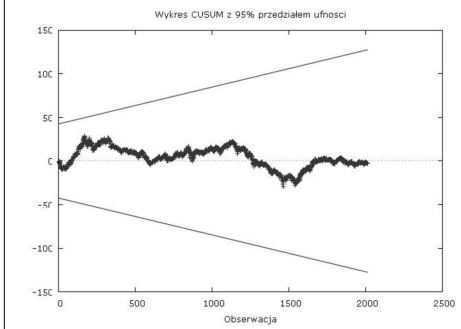
Fundusz Novo model T-M (15)



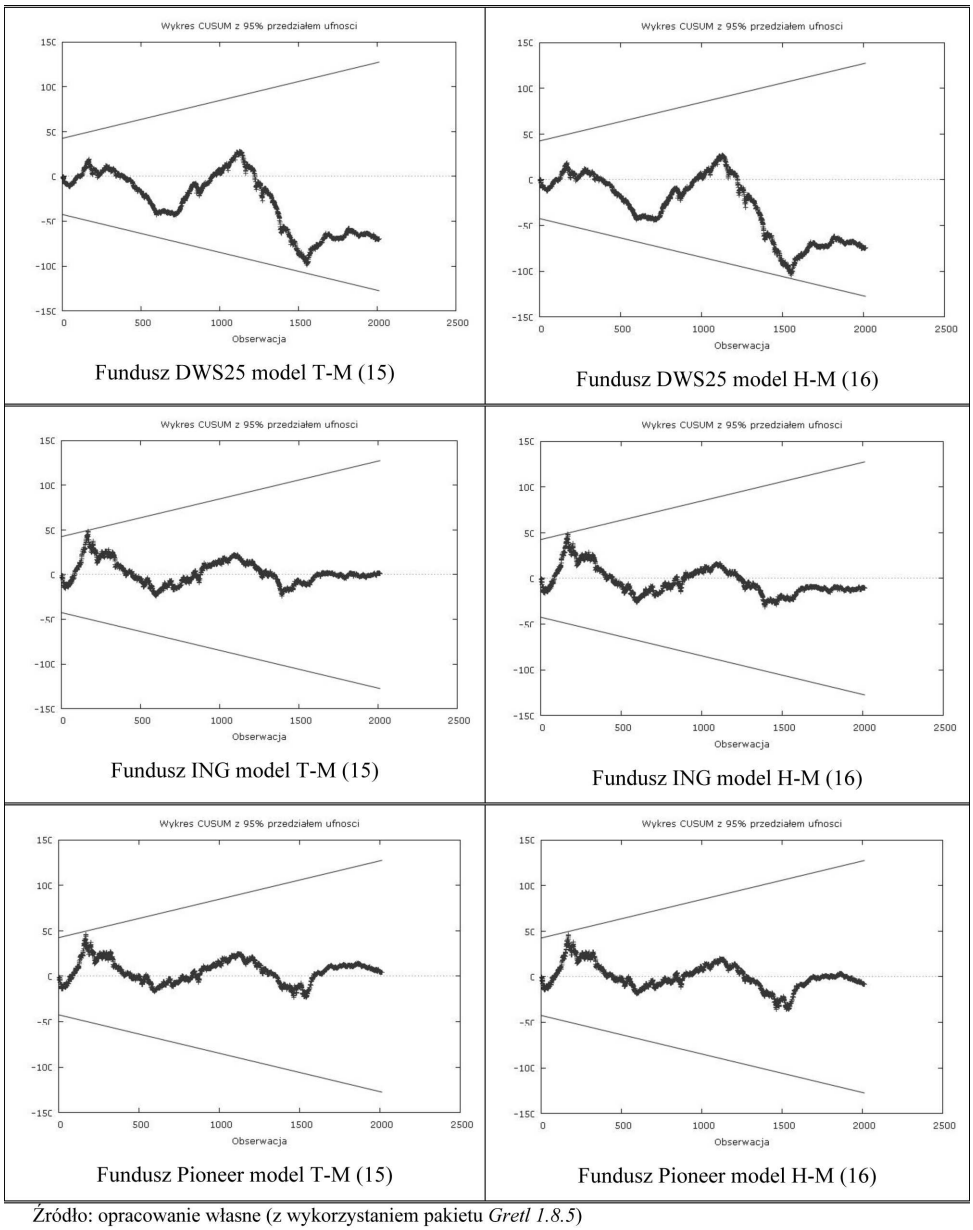
Fundusz Novo model H-M (16)



Fundusz UniKorona model T-M (15)



Fundusz UniKorona model H-M (16)



Źródło: opracowanie własne (z wykorzystaniem pakietu *Gretl 1.8.5*)

Rysunek 1. Wykresy testu *CUSUM* stabilności parametrów modeli market-timing wybranych funduszy akcji polskich w okresie 2.01.2003r.-31.12.2010r.

Źródło: opracowanie własne (z wykorzystaniem pakietu Gretl 1.8.5)

7. PODSUMOWANIE

Dychotomiczny podział umiejętności zarządzającego portfelem, zaproponowany przez Famę [10], nie jest możliwy w praktyce, ponieważ selektywność papierów wartościowych opiera się nie tylko na analizie fundamentalnej sytuacji emitentów, ale również na porównaniu wyników tej analizy z bieżącą ceną rynkową. Klasyczny współczynnik alfa Jensena w modelach market-timing, jako miara umiejętności w zakresie selektywności aktywów, nie powinien być rozpatrywany w oderwaniu od interpretacji pozostałych parametrów modelu. Jeśli wartość $\hat{\alpha}_P$ jest dodatnia, można stwierdzić, że zarządzający najprawdopodobniej posiada lepszą umiejętność selekcji aktywów niż ogół inwestorów, nie można natomiast powiedzieć, iż jest to zasługa wyłącznie tej umiejętności, ponieważ za część ponadprzeciętnej stopy zwrotu może odpowiadać zmienność w czasie współczynnika β_P [8]. Zgodnie z interpretacją Jensena [19], dodatnia wartość estymatora parametru α_P może też wynikać z dodatniego obciążenia tego estymatora.

Politechnika Białostocka

LITERATURA

[1] Admati A., Bhattacharya S., Pfleiderer P., Ross S., [1986], *On timing and selectivity*, „The Journal of Finance", 41 (July), str. 715-73.
[2] Bollen N. P. B., Busse J. A., [2001], *On the timing ability of mutual fund managers*, „The Journal of Finance", Vol. LVI, No. 3, str. 1075-1094.
[3] Brown R.L., Durbin J., Evans J.M. [1975], *Techniques for testing the constancy of regression relationships over time*, „Journal of Royal Statistical Society", 37, No. 2, str. 149-192.
[4] Chang E., Lewellen W., [1984], *Market timing and mutual fund investment performance,* ..Journal of Business" 57, str. 57-72.
[5] Chen C.R., Stockum S., [1986], *Selectivity, market timing and random beta behavior of mutual funds: a generalized model*, „The Journal of Financial Research" 9, No. 1, str. 87-96.
[6] Cheng-few Lee, Rahman S., [1990], *Market timing, selectivity, and mutual fund performance: an empirical investigation*, „Journal of Business" 63, No. 2, str. 261-276.
[7] Connor G., Korajczyk R.A., [1991], *The attributes, behaviour and performance of US mutual funds,* „Review of Quantitative Finance and Accounting", Vol. 1, str.5-26.
[8] Czekaj J., Woś M., Żarnowski J, [2001], *Efektywność giełdowego rynku akcji w Polsce. Z perspektywy dziesięciolecia*, PWN, Warszawa.
[9] Fama E., [1968], *Risk, return and equilibrium: some clarifying comments,* „The Journal of Finance" 23, str. 29-40.
[10] Fama E., [1972], *Components of investment performance*, „The Journal of Finance" 27, No. 3, str. 551-567.
[11] Ferson W.E., Schadt R.W., [1996], *Measuring fund strategy and performance in changing economic conditions,* „The Journal of Finance" 51, No. 2, June, str. 425-461.
[12] Fletcher J., [1995], *An examination of the selectivity and market timing performance of UK unit trusts,* „Journal of Business Finance & Accounting" 22, str. 143-156.
[13] Goetzmann W.N., Ingersoll J. Jr., Ivković Z., [2000], *Monthly measurement of daily timers,* „Journal of Financial and Quantitative Analysis", Vol. 35, No. 3, September, str. 257-290.

[14] Grant D., [1977], *Portfolio performance and the „cost" of timing decisions,* „The Journal of Finance" 32, str. 837-846.
[15] Grinblatt M., Titman S., [1994], *A study of monthly mutual fund returns and performance evaluation techniques*, „Journal of Financial and Quantitative Analysis", 29, str. 419-444.
[16] Henriksson R., Merton R., [1981], *On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills*, „Journal of Business" 54, No. 4, str.513-533.
[17] Henriksson R., [1984], *Market timing and mutual fund performance: an empirical investigation*, „Journal of Business" 57, str. 73-96.
[18] Jensen M.C., [1972], *Optimal utilization of market forecasts and the evaluation of investment performance,* [in:] Szego G.P., Shell K. (eds.) „Mathematical Methods in Investment and Finance". Amsterdam.
[19] Jensen M.C., [1968], *The performance of mutual funds in the period 1945-1964*, „The Journal of Finance" 23, str. 389-416.
[20] Kao G., Cheng L., Chan K., [1998], *International mutual fund selectivity and market timing during up and down market conditions,* „The Financial Review", 33, str. 127-144.
[21] Kufel T., [2007], *Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu Gretl*, PWN, Warszawa.
[22] Merton R., [1981], *On market timing and investment performance. I. An equilibrium theory of value for market forecasts,* „Journal of Business" 54, No. 3, str. 363-406.
[23] Olbryś J. [2009], *Conditional market-timing models for mutual fund performance evaluation,* „Prace i Materiały Wydziału Zarządzania Uniwersytetu Gdańskiego"4/2, str. 519-532.
[24] Olbryś J., [2008], *Parametric tests for timing and selectivity in Polish mutual fund performance*, „Optimum. Studia Ekonomiczne", Wydawnictwo Uniwersytetu w Białymstoku, 3(39)/2008, str.107-118.
[25] Olbryś J., [2008], *Parametryczne testy umiejętności wyczucia rynku – porównanie wybranych metod na przykładzie OFI akcji,* [w:] Z. Binderman (red.) „Metody ilościowe w badaniach ekonomicznych IX", Wydawnictwo SGGW w Warszawie, str. 81-88
[26] Olbryś J., [2008], *Ocena umiejętności stosowania strategii market-timing przez zarządzających portfelami funduszy inwestycyjnych a częstotliwość danych,* „Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania" Nr 10, Uniwersytet Szczeciński, Szczecin, str. 96-105.
[27] Olbryś J., Karpio A., [2009], *Market-timing and selectivity abilities of Polish open-end mutual funds managers*, [w:] P.Chrzan, T.Czernik (red.) „Metody matematyczne, ekonometryczne i komputerowe w finansach i ubezpieczeniach 2007, Wydawnictwo AE im. K. Adamieckiego w Katowicach, str. 437-443.
[28] Prather L.J., Middleton K.L., [2006], *Timing and selectivity of mutual fund managers: An empirical test of the behavioral decision-making theory*, „Journal of Empirical Finance", 13, str. 249-273.
[29] Rao S., [2000], *Market timing and mutual fund performance,* „American Business Review", 18, str. 75-79.
[30] Rao S., [2001], *Mutual fund performance during up and down market conditions,* „Review of Business", 22, str. 62-65.
[31] Romacho J. C., Cortez M. C., [2006], *Timing and selectivity in Portuguese mutual fund performance*, „Research in International Business and Finance" 20, str. 348-368.
[32] Sharpe W.F., Alexander G.J., Bailey J.V., [1999], *Investments,* Prentice Hall, New Jersey.
[33] Sharpe W.F., [1966], *Mutual fund performance*, „Journal of Business" 39, str. 119-138.
[34] Sharpe W.F., [1992], *Asset allocation: Management style and performance measurement*, „The Journal of Portfolio Management" Winter, str. 7-19.
[35] Stein C.M., [1981], *Estimation of the mean of a multivariate normal distribution,* „The Annals of Statistics" 9, No. 6, str. 1135-1151.
[36] Stevens G.V.G., [1971], *Two problems in portfolio analysis: conditional and multiplicative random variables,* „Journal of Financial and Quantitative Analysis" 6 (December), str. 1235-1250.

[37] Treynor J., [1965], *How to rate management of investment funds*, „Harvard Business Review" 43, str. 63-75.
[38] Treynor J., Mazuy K., [1966], *Can mutual funds outguess the market?,* „Harvard Business Review", 44, str. 131-136.

OBCIĄŻENIE ESTYMATORA WSPÓŁCZYNNIKA ALFA JENSENA A INTERPRETACJE PARAMETRÓW KLASYCZNYCH MODELI MARKET-TIMING

Streszczenie

W 1968 roku ukazał się w „The Journal of Finance" artykuł M.C. Jensena *The performance of mutual funds in the period 1945 – 1964*, w którym m.in. wprowadzony został współczynnik α, zwany w literaturze przedmiotu alfą Jensena, jako miara umiejętności zarządzającego portfelem inwestycyjnym w zakresie selektywności aktywów. Wnioski Jensena odnośnie interpretacji można podsumować następującym stwierdzeniem: estymator $\hat{\alpha}$ może być dodatni z dwóch powodów: 1) dodatkowych zysków osiąganych przez menadżera wykorzystującego swoje umiejętności w zakresie doboru aktywów; 2) dodatniego obciążenia estymatora parametru α. W roku 1977, również w „The Journal of Finance", ukazał się z kolei artykuł D. Granta *Portfolio performance and the „cost" of timing decisions*, w którym Autor podważa prawdziwość wzoru przedstawionego przez Jensena pisząc wprost o błędzie matematycznym w przekształceniach. Prezentuje własne wyprowadzenie wzoru i podaje nową interpretację, według której estymator współczynnika alfa jest obciążony ujemnie. Wprowadza to sporo zamieszania w literaturze przedmiotu, ponieważ część autorów powołuje się w swoich pracach na interpretację Jensena, natomiast część na interpretację Granta. W konsekwencji wnioski dotyczące wyników badań empirycznych na rynkach funduszy inwestycyjnych, w zakresie umiejętności stosowania przez zarządzających funduszami technik market-timing oraz selektywności aktywów, nie są jednoznaczne.

Głównym punktem artykułu jest wyprowadzenie spornego wzoru, w celu wyciągnięcia ostatecznych wniosków dotyczących statystycznego oddziaływania obciążenia estymatora współczynnika alfa Jensena na pozostałe parametry modelu regresji. Dodatkowo przedstawione zostaną przykłady klasycznych modeli market-timing polskich funduszy inwestycyjnych akcji, jak również zostanie zbadana stabilność parametrów otrzymanych modeli ekonometrycznych w okresie styczeń 2003 – grudzień 2010.

Słowa kluczowe: alfa Jensena, selektywność, modele market-timing

THE INFLUENCE OF THE BIAS OF THE JENSEN'S ALPHA COEFFICIENT ESTIMATOR ON INTERPRETING THE PARAMETERS OF THE CLASSICAL MARKET-TIMING MODELS

Summary

Jensen's alpha coefficient was introduced in 1968 as a measure of the manager's ability to forecast future security prices (M.C. Jensen, *The performance of mutual funds in the period 1945 – 1964*, „The Journal of Finance"). The paper concluded that the performance measure α may be positive for two reasons: 1) extra returns earned on the portfolio due to the manager's ability and/or 2) a positive $\hat{\alpha}$ estimator bias.

In 1977, D.Grant (*Portfolio performance and the „cost" of timing decisions*, „The Journal of Finance") stated that Jensen's original work contained a mathematical error and a conceptual problem. He propounded a new equation which differed from that of Jensen in both direction and magnitude of the $\hat{\alpha}$ estimator bias. Grant's new approach caused serious confusion about the subject as some authors now quote after Jensen and some after Grant. Consequently, the conclusions concerning managed mutual funds' portfolios performance are often ambiguous. The aim of this paper is to resolve this ambiguity and to determine which alpha coefficient interpretation is correct. The paper contains a proof of Jensen's equation. We also present the examples of the classical T-M and H-M market-timing models in the case of Polish equity open-end mutual funds and test their stability in the period Jan 2003-Dec 2010.

Keywords: Jensen's alpha coefficient, selectivity, classical market-timing models

PRZEGLĄD STATYSTYCZNY
R. LVIII – ZESZYT 1-2 – 2011

JOANNA KISIELIŃSKA

DOKŁADNA METODA BOOTSTRAPOWA I JEJ ZASTOSOWANIE DO ESTYMACJI WARIANCJI

1. WPROWADZENIE

Dana jest zmienna losowa X o nieznanym rozkładzie F. Interesuje nas parametr rozkładu, który oznaczymy jako θ. Jeśli parametru nie można wyznaczyć bezpośrednio, konieczne jest pobranie próby losowej oraz dobranie odpowiedniego estymatora parametru θ. Estymator jest statystyką określoną na przestrzeni prób. Próbę losową oznaczymy jako $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, jej realizację jako $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, zaś estymator parametru θ jako $\hat{\theta} = t(\mathbf{X})$.

Efron [4] zaproponował metodę, którą nazwał bootstrap polegającą na losowaniu z uzyskanej próby (próby pierwotnej) $\mathbf{x}$, kolejnych prób wtórnych (prób bootstrapowych) o liczebności n. Losowanie odbywa się ze zwracaniem przy założeniu jednakowych prawdopodobieństw równych $1/n$ wylosowania, każdej z wartości x_i, dla i=1, …, n. Generowany jest w ten sposób rozkład $\hat{F}$, zwany rozkładem bootstrapowym z próby.

Próba bootstrapowa oznaczana jest jako $\mathbf{X}^* = (X_1^*, X_2^*, \cdots, X_n^*)$, a dowolna jej realizacja jako $\mathbf{x}^* = (x_1^*, x_2^*, \cdots, x_n^*)$. Statystyka $\hat{\theta}$ dla próby bootstrapowej oznaczana jest jako $\hat{\theta}* = t(\mathbf{X}*)$.

Istotą metody jest aproksymacja rozkładu statystyki $\hat{\theta}$, rozkładem statystyki bootstrapowej $\hat{\theta}*$. Efron [4] zaproponował trzy metody wyznaczania rozkładu $\hat{\theta}*$:

1. przeprowadzenie właściwych rozważań teoretycznych,
2. zastosowanie aproksymacji Monte Carlo,
3. aproksymację rozkładu $\hat{\theta}*$ poprzez rozwinięcie funkcji t w szereg Taylora.

Jeśli wybrane zostanie rozwiązanie 2 konieczne jest określenie liczby losowanych prób bootstrapowych N. Efron [5] opierając się na bootstrapowej wariancji stwierdza, że wystarczy niewielka liczba losowań, aby otrzymać wystarczającą dokładność. Z taką oceną nie zgadzają się Booth i Sarkar [1], którzy zastosowali aproksymację rozkładu względnej wariancji bootstrapowej rozkładem normalnym. Pozwoliło to na oszacowanie N dla określonego poziomu błędu na założonym poziomie ufności. Okazało się, że uzyskanie błędu poniżej 10% przy poziomie ufności 0,95 wymaga przyjęcia N około 800. Domański i Pruska [3] (str. 261) stwierdzają, że przyjmuje się $N \geqslant 1000$.

Prowadząc rozważania nad metodą bootstrapową zadać można sobie pytanie, czy losowanie próby bootstrapowej $\mathbf{X}^*$, z wcześniej już uzyskanej próby pierwotnej $\mathbf{X}$ jest konieczne? Losowanie próby jest niezbędne, jeśli zbadanie całej populacji nie jest

możliwe lub jest zbyt kosztowne. Posługiwanie się próbą zamiast populacją ma swoje istotne implikacje będące w obszarze zainteresowań statystyki matematycznej.

Zwróćmy uwagę, że podstawową własnością próby jest jej skończony rozmiar. Określony dla niej rozkład bootstrapowy $\hat{F}$ jest prostym rozkładem dyskretnym. Rozkład dowolniej statystyki określonej dla n dyskretnych zmiennych losowych o skończonej liczbie realizacji, nie musi być szacowany, ponieważ może być po prostu wyznaczony. Kwestią otwartą pozostaje jedynie, jak dużego nakładu obliczeń wymaga takie podejście, o czym mowa będzie dalej.

Rozkład bootstrapowy jest w istocie rozkładem empirycznym określonym dystrybuantą[1)]:

$$F_n(t) = \frac{\#\ \{1 \leqslant i \leqslant n : X_i \leqslant t\}}{n} \text{ dla } t \in R^1 \tag{1}$$

gdzie $X_1, X_2, \ldots, X_n$ jest ciągiem ciągłych i niezależnych zmiennych losowych o jednakowych rozkładach określonych dystrybuantą F.

Z twierdzenia Gliwienki-Cantelliego (podstawowego twierdzenia statystyki matematycznej) wynika, że:

$$D_n = \sup_{-\infty<x<+\infty} |F_n(x) - F(x)| \tag{2}$$

dąży do 0 z prawdopodobieństwem 1.

Twierdzenie to upoważnia do wnioskowania statystycznego – formułowania stwierdzeń dotyczących populacji na podstawie próby.

2. DOKŁADNA METODA BOOTSTRAPOWA

Załóżmy, że z populacji opisanej zmienną losową X wylosowano n elementową próbę pierwotną $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Ponieważ dla niektórych $i \neq j$ może zachodzić $x_i = x_j$, zredukujmy[2)] rozmiar wylosowanej próby do k różnych wartości. Prawdopodobieństwa p_i wylosowania z próby realizacji x_i, gdzie i=1, 2,..., k nie muszą być wówczas jednakowe (jak to ma miejsce dla próby bootstrapowej).

W próbie **x**, każda wartość x_i jest realizacją zmiennej losowej X o rozkładzie F. Zgodnie z twierdzeniem Gliwienki-Cantelliego, rozkład F może być przybliżony rozkładem empirycznym F_n.

Wprowadźmy pojecie dyskretnej zmiennej losowej próby, którą oznaczyć można jako X^D. Zbiorem realizacji tej zmiennej jest pierwotna próba losowa $\{x_1, x_2, \ldots, x_k\}$. Poszczególne wartości przyjmowane są z prawdopodobieństwami $p_i^{\mathrm{D}} = P\left(X^{\mathrm{D}} = x_i\right) = p_i$[3)], dla i= 1, 2, ... , k. Rozkład ten jest równoważny rozkładowi empirycznemu

[1)] Zapis zgodnie z Zieliński [9] str. 11, gdzie # oznacza liczebność.

[2)] Redukcja ta nie jest konieczna, lecz wskazana, ponieważ pozwala na zmniejszenie wymiaru problemu.

[3)] Prawdopodobieństwa te byłyby jednakowe i równe $1/n$ gdyby nie przeprowadzono redukcji wymiaru próby do k różnych wartości.

i oznaczony zostanie jako F^D. W takim przypadku n elementową wtórną próbę losową można zapisać jako $\mathbf{X}^{\mathrm{D}} = \left(X_1^{\mathrm{D}}, X_2^{\mathrm{D}}, \cdots, X_n^{\mathrm{D}}\right)$, a dowolną jej realizację jako $\mathbf{x}^{\mathrm{D}} = \left(x_1^{\mathrm{D}}, x_2^{\mathrm{D}}, \cdots, x_n^{\mathrm{D}}\right)$. Zmienne X_i^{D}, dla i= 1, 2, … , n mają rozkłady F^D. Statystykę $\hat{\theta}*$ dla próby bootstrapowej oznaczyć można wówczas jako $\hat{\theta}^{\mathrm{D}} = t\left(\mathbf{X}^{\mathrm{D}}\right)$. Rozkład statystyki $\hat{\theta}$ przybliżać będziemy rozkładem statystyki $\hat{\theta}^{\mathrm{D}}$. Zauważmy, że problemy związane z ewentualnym obciążeniem, zgodnością i efektywnością estymatora dotyczą estymatora $\hat{\theta}$. W dokładnej metodzie bootstrapowej wartość estymatora $\hat{\theta}^{\mathrm{D}}$ jest obliczana, a nie szacowana na podstawie próby. Metoda nie wprowadza więc dodatkowego obciążenia.

Dla pojedynczej b-tej realizacji wtórnej próby $\mathbf{x}^{\mathrm{D}b} = \left(x_1^{\mathrm{D}b}, x_2^{\mathrm{D}b}, \cdots, x_n^{\mathrm{D}b}\right)$ należy obliczyć wartość $\hat{\theta}^{\mathrm{D}}(b) = t\left(\mathbf{x}^{\mathrm{D}b}\right) = t\left(x_1^{\mathrm{D}b}, x_2^{\mathrm{D}b}, \cdots, x_n^{\mathrm{D}b}\right)$ oraz prawdopodobieństwo jej wylosowania. Ze względu na redukcję wymiaru n elementowej próby do k różnych wartości, prawdopodobieństwa wylosowania poszczególnych prób bootstrapowych nie są już jednakowe. Prawdopodobieństwo wylosowania b-tej próby jest równe $p^{Db} = P\left(\hat{\theta}^{\mathrm{D}} = \hat{\theta}^{\mathrm{D}}(b)\right) = \prod_{i=1}^{n} p_i^{\mathrm{D}b}$, gdzie $p_i^{\mathrm{D}b} = P\left(X^{\mathrm{D}} = x_i^{\mathrm{D}b}\right)$. Liczba możliwych realizacji wtórnej próby jest równa $B = k^n$. Poprawnie napisany algorytm powinien zapewnić spełnienie warunku: $\sum_{b=1}^{B} p^{\mathrm{D}b} = 1$.

Mając wyznaczone wartości $\hat{\theta}^{\mathrm{D}}(b)$ oraz prawdopodobieństwa p^{Db}, dla b=1, 2,. . . , B można określić rozkład estymatora $\hat{\theta}^{\mathrm{D}}$. Estymator ten, będąc funkcją dyskretnych zmiennych losowych o jednakowym rozkładzie F^D, jest również zmienną losową dyskretną. Rozkład estymatora pozwala budować przedziały ufności, czy testować hipotezy statystyczne.

Jeżeli znany jest rozkład estymatora można wyznaczyć jego wartość oczekiwaną i odchylenie standardowe. Wartość oczekiwana jest bootstrapowym oszacowaniem parametru θ, zaś odchylenie standardowe błędem tego szacunku. Wartość oczekiwana i odchylenie standardowe obliczyć należy tak jak dla zmiennej dyskretnej. Oznaczając ocenę parametru jako $\hat{\theta}^D(\cdot)$, a błąd jako $s_{\hat{\theta}^{\mathrm{D}}}^{\mathrm{D}}$ [4)] otrzymujemy:

$$\hat{\theta}^{\mathrm{D}}(\cdot) = \sum_{b=1}^{B} \hat{\theta}^{\mathrm{D}}(b) \cdot p^{\mathrm{D}b}, \quad s_{\hat{\theta}^{\mathrm{D}}}^{\mathrm{D}} = \sqrt{\sum_{b=1}^{B} \left(\hat{\theta}^{\mathrm{D}}(b) - \hat{\theta}^{\mathrm{D}}(\cdot)\right)^2 \cdot p^{\mathrm{D}b}} \tag{3}$$

Chcąc zastosować metodę dokładnego bootstrapu warto zastanowić się, jak ma się liczba wszystkich możliwych prób wtórnych B wobec zalecanej liczby prób wtórnych w klasycznej metodzie bootstrapowej. Np. dla k=10 i n=20 otrzymujemy $B = 10^{20}$,

[4)] Formuła (3) obowiązuje gdy brane są pod uwagę wszystkie próby bootstrapowe. Jeśli rozważany jest jedynie ich podzbiór należałoby dokonać korekty o czynnik $\sqrt{\frac{n}{n-1}}$.

co jest liczbą bardzo dużą, znacznie większą niż przykładowe N=1000. Jak pokazane zostanie dalej tak dużą liczbę powtórzeń można aktualnie wygenerować. Pionierska praca Efrona pochodzi z roku 1979. W tamtym okresie wykonanie tej liczby obliczeń w sensownym czasie było niemożliwe. Ponieważ nie można było zbadać całej „populacji" określonej próbą pierwotną, konieczne było pobieranie z „próby traktowanej jako populację" kolejnych prób – prób wtórnych.

W metodzie bootstrapowej ciąg uzyskanych wartości estymatora porządkowany jest w kolejności od najmniejszego do największego, co pozwala np. wyznaczyć granice przedziałów ufności metodą percentyli (Efron, Tibshirani [6]). W dokładnej metodzie bootstrapowej teoretycznie można również tak postąpić. Liczba możliwych realizacji statystyki $\hat{\theta}^{\mathrm{D}}$ jest wprawdzie bardzo duża, ale po pierwsze, część realizacji z pewnością będzie się powtarzać, a po drugie i tak wskazane jest pogrupowanie rezultatów w histogramie. Utworzenie histogramu będzie dla dużych problemów niezbędne[5)], ale wiązać się będzie z utratą pewnej informacji. Podkreślić należy, że mimo tego dokładną metodą bootstrapową można uzyskać bardzo dokładne oszacowanie granic przedziałów ufności, ponieważ szerokość przedziałów w histogramie nie musi być jednakowa. W zakresach wymagających podania dokładnych prawdopodobieństw (czy dystrybuanty), szerokość przedziału może być bardzo, niemal dowolnie mała.

Najprostszym sposobem wygenerowania wszystkich wtórnych prób dla rozkładu dyskretnego będzie rekurencyjne pobieranie kolejnych elementów z próby. Jeśli wystąpiły powtórzenia, zbiór z którego pobierane są wartości redukuje się do wymiaru k. Tak skonstruowany algorytm zaliczyć można do kategorii brute force. Algorytmy tego typu są uznawane za nieefektywne.

Liczbę generowanych realizacji prób bootstrapowych można zredukować, ponieważ w próbie wtórnej niektóre wartości będą się powtarzać – losowanie odbywa się z powtórzeniami. Problem taki dla przypadku gdy prawdopodobieństwo wylosowania każdego elementu próby jest jednakowe i elementy są jednakowe, przedstawiony jest w książce Fellera [7] w rozdziale II.5 str. 38[6)]. Dalsze rozważania prowadzone będą dla przypadku, gdy prawdopodobieństwa te nie są jednakowe, a losowane elementy mogą być różne.

Każdą n elementową wtórną b-tą próbę bootstrapową losowaną ze zbioru k elementowego można zapisać jako:

$$\mathbf{x}^{\mathrm{D}b} = (a_{1b} \times x_1, a_{2b} \times x_2, \cdots, a_{kb} \times x_k) \tag{4}$$

gdzie każde $a_{ib} \geqslant 0$, dla i= 1, 2, …, k jest liczbą wystąpień w próbie b-tej i-tego elementu próby pierwotnej. Liczby te muszą spełniać warunek:

$$\sum_{i=1}^{k} a_{ib} = n \tag{5}$$

[5)] Przykładowo liczba realizacji 10^{20} wymaga ponad 80 eksbibajtów komórek pamięci.

[6)] Jak zauważa Booth [2] w prezentacji elektronicznej: Monte Carlo and the boothstrap.

przy czym część z nich może być równa 0. Jeżeli dla wybranego i zachodzi a_{ib}=0, oznacza to, że element x_i nie wystąpił w b-tej próbie wtórnej.

Prawdopodobieństwo wylosowania pojedynczej próby określonej (4) jest równe:

$$p^{\mathrm{D}b} = \prod_{i=1}^{k} \left(p_i^{\mathrm{D}b}\right)^{a_{ib}} \tag{6}$$

Pozostaje jeszcze określenie ile prób można wygenerować dla danego ciągu liczb $a_{ib} \geqslant 0$, dla i= 1, 2, ..., k, spełniających warunek (5). Trzeba uwzględnić wszystkie m_b elementowe kombinacje, gdzie $m_b = \#\{a_{ib} \neq 0 : i = 1, 2, ..., k\}$ ze zbioru k elementowego. Kombinacje należy następnie permutować, ale jedynie na pozycjach, dla których współczynniki a_{ib} są unikalne.

3. ROZKŁAD GRANICZNY ESTYMATORA WARIANCJI Z PRÓBY BOOTSTRAPOWEJ

Dokładna metoda bootstrapowa wykorzystana zostanie do szacowania wariancji pewnej zmiennej losowej o nieznanym rozkładzie. Weryfikację poprawności metody można przeprowadzić porównując rozkład estymatora wariancji z rozkładem granicznym, który stosować można gdy próba jest duża ($n \geqslant 30$).

Pobierając n elementową próbę losową określamy dyskretną zmienną losową próby X^D o zbiorze realizacji $\{x_1, x_2, \ldots, x_k\}$ i rozkładzie prawdopodobieństwa określonym wartościami $p_i = \mathrm{P}(X^D = x_i)$, dla i= 1,2, ... , k, przy czym $\sum_{i=1}^{k} p_i = 1$.

Wartość oczekiwana μ^D, odchylenie standardowe σ^{D} oraz czwarty moment centralny μ_4^{D} zmiennej X^D są równe odpowiednio:

$$\mu^{\mathrm{D}} = \sum_{i=1}^{k} x_i \cdot p_i, \quad \sigma^{\mathrm{D}} = \sqrt{\sum_{i=1}^{k} (x_i - \mu^{\mathrm{D}})^2 \cdot p_i}, \quad \mu_4^{\mathrm{D}} = \sum_{i=1}^{k} \left(x_i - \mu^{\mathrm{D}}\right)^4 \cdot p_i \tag{7}$$

Wiadomo, że rozkład wariancji z próby $S^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2$ pobranej z populacji o rozkładzie normalnym N(μ,σ) jest rozkładem asymptotycznie normalnym $\mathrm{N}\left(\frac{n-1}{n} \cdot \sigma^2, \sqrt{2 \cdot \frac{\sigma^4}{n}}\right)$.

Rozkład graniczny wariancji z próby S^2 pobranej z populacji o dowolnym rozkładzie z parametrami μ i σ będzie też rozkładem normalnym (wariancja jest uśrednionym

kwadratem odchyleń od wartości przeciętnych). Wartość oczekiwana rozkładu granicznego oraz jego wariancja są określone wzorem[7)]:

$$\mathrm{E}S^2 = \frac{n-1}{n}\cdot\sigma^2, \quad \mathrm{D}^2S^2 = \frac{\mu_4-\sigma^4}{n} - \frac{2\cdot\left(\mu_4-2\cdot\sigma^4\right)}{n^2} + \frac{\mu_4-3\cdot\sigma^4}{n^3} \tag{8}$$

gdzie: μ_4 jest momentem centralnym rzędu czwartego rozpatrywanego rozkładu.

Rozkłady, o których mowa powyżej są rozkładami granicznymi dla estymatora S^2, a nie estymatora $\hat{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$. Ponieważ $\hat{S}^2 = \frac{n}{n-1}S^2$, należy dokonać korekty parametrów rozkładów granicznych. Wartość oczekiwaną estymatora S^2 należy pomnożyć[8)] przez $\frac{n}{n-1}$, a wariancję przez $\left(\frac{n}{n-1}\right)^2$.

Rozkłady graniczne nieobciążonego estymatora wariancji oznaczone zostaną jako GV1 (rozkład graniczny nie wymagający normalności rozkładu zmiennej losowej w populacji) i GV2 (wymagający normalności rozkładu) są więc następujące:

$$\mathrm{GV1}: \mathrm{N}\left(\left(\sigma^{\mathrm{D}}\right)^2, \frac{n}{n-1}\cdot\sqrt{\frac{\mu_4^{\mathrm{D}}-\left(\sigma^{\mathrm{D}}\right)^4}{n} - \frac{2\cdot\left(\mu_4^{\mathrm{D}}-2\cdot\left(\sigma^{\mathrm{D}}\right)^4\right)}{n^2} + \frac{\mu_4^{\mathrm{D}}-3\cdot\left(\sigma^{\mathrm{D}}\right)^4}{n^3}}\right)$$

$$\mathrm{GV2}: \mathrm{N}\left(\left(\sigma^{\mathrm{D}}\right)^2, \frac{n}{n-1}\cdot\sqrt{\frac{2\cdot\left(\sigma^{\mathrm{D}}\right)^4}{n}}\right) \tag{9}$$

4. WYNIKI

Zakładamy, że dana jest zmienna losowa X o nieznanym rozkładzie, którego wariancję chcemy oszacować. Pobieramy n elementową próbę losową i zakładamy postać estymatora wariancji $\hat{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$. Przykładowy rozkład empiryczny próby przedstawiony jest w tabeli 1. Niech rozkład ten reprezentuje zmienna losowa X^D. Wartość oczekiwana i wariancja zmiennej X^D są odpowiednio równe μ^D= 5,174 oraz $\left(\sigma^{\mathrm{D}}\right)^2$= 3,989724. Rozkład estymatora wariancji zmiennej losowej X przybliżamy rozkładem statystyki bootstrapowej. Pobranie próby losowej można interpretować jako dyskretyzację ciągłej zmiennej X. Stosując metodę bootstrapową rozkład estymatora pewnego parametru rozkładu zmiennej X, przybliżamy rozkładem estymatora dyskretnej zmiennej X^D reprezentującej próbę. Różnica miedzy klasyczną metodą, a dokładną

[7)] Smirnow, Dunin-Barkowski [8] str. 237.

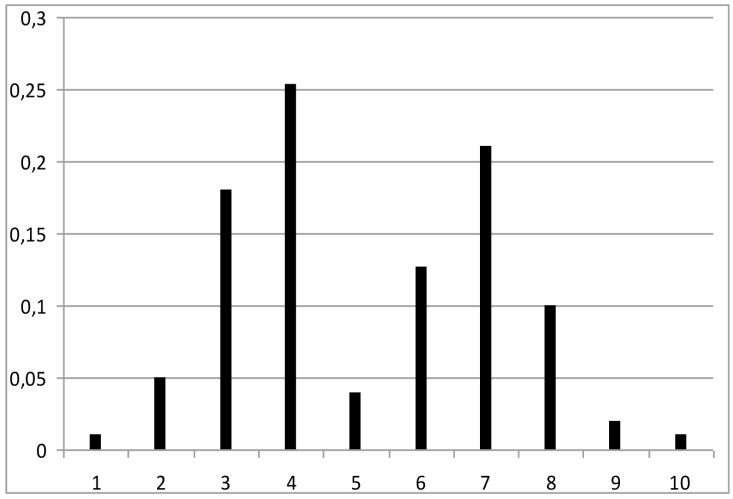
[8)] Jeżeli n jest bardzo duże iloraz $\frac{n}{n-1}$ jest praktycznie równy jeden. Dla próby 30 elementowej (a dla tej liczebności próby można już stosować rozkłady graniczne) jest równy 1,0345, zaś jego kwadrat 1,0702 – o tyle więc należałoby skorygować wariancję.

polega jedynie na liczbie wziętych pod uwagę prób wtórnych. W metodzie klasycznej losowany jest pewien podzbiór możliwych prób wtórnych, w metodzie dokładnej zaś brane są pod uwagę wszystkie próby wtórne. W klasycznej metodzie bootstrapowej zamiast populacji prób wtórnych badamy więc jedynie pewien jej podzbiór. W dokładnej metodzie bootstrapowej bierzemy pod uwagę całą populację prób wtórnych. Mając do dyspozycji całą populację rozkład, bądź jego parametry nie muszą być szacowane, ponieważ mogą być obliczane – problem na tym etapie jest zagadnieniem z zakresu statystyki opisowej, a nie matematycznej.

Tabela 1

Rozkład zmiennej losowej X^D reprezentującej próbę pierwotną

x_i	p_i
1	0,010
2	0,050
3	0,180
4	0,253
5	0,040
6	0,127
7	0,210
8	0,100
9	0,020
10	0,010



Źródło: Opracowanie własne

W tabeli 2 przedstawiono stablicowany rozkład estymatora wariancji zmiennej losowej X^D (stanowiący przybliżenie rozkładu estymatora wariancji zmiennej ciągłej X), wyznaczony metodą dokładnego bootstrapu i oznaczony jako DBV, oraz dwa normalne rozkłady graniczne dla wariancji GV1 i GV2 określone wzorem (9).

Rozkłady podano w dwóch wariantach. W pierwszym założono n=20, w drugim n=30. Podkreślić należy, że w przypadku stosowania metod bootstrapowych wartość n wynika bezpośrednio z liczebności próby. Założenie dwóch wartości dla n traktować należy jedynie jako eksperyment symulacyjny mający na celu zbadanie ewentualnych podobieństw i różnic między faktycznym rozkładem estymatora a rozkładami granicznymi.

Pewnego komentarza wymaga jeszcze sposób doboru szerokości przedziału tablicowania. W ogólnym przypadku może być ona dowolnie mała. Dla rozkładów granicznych (ciągłych), dla każdego dowolnie małego przedziału istnieje większe od 0

Tabela 2

Rozkład estymatora wariancji zmiennej losowej X^D uzyskany dokładną metodą bootstrapową oraz rozkłady graniczne

	n=20			*n*=30		
	DBV	GV1	GV2	DBV	GV1	GV2
Średnia	3,98972	3,98972	3,98972	3,98972	3,98972	3,98972
Odchylenie st.	0,91790	0,91790	1,32806	0,73650	0,73650	1,06566
Przedziały	DBV	GV1	GV2	DBV	GV1	GV2
<0,00; 0,25)	2,77E-08	2,31E-05	0,002432	8,73E-12	1,91E-07	0,000225
<0,25; 0,50)	9,47E-07	4,87E-05	0,001867	1,52E-09	8,87E-07	0,000304
<0,50; 0,75)	5,73E-06	1,36E-04	0,003057	3,14E-08	4,36E-06	0,000654
<0,75; 1,00)	2,91E-05	0,000355	0,004832	3,66E-07	1,92E-05	0,001329
<1,00; 1,25)	9,40E-05	0,000856	0,007372	2,84E-06	7,50E-05	0,002560
<1,25; 1,50)	0,000361	0,001921	0,010858	2,03E-05	0,000262	0,004666
<1,50; 1,75)	0,001463	0,004003	0,015437	0,000110	0,000817	0,008051
<1,75; 2,00)	0,003616	0,007749	0,021185	0,000607	0,002272	0,013153
<2,00; 2,25)	0,009544	0,013933	0,028064	0,002482	0,005634	0,020342
<2,25; 2,50)	0,021507	0,023274	0,035886	0,008270	0,012467	0,029782
<2,50; 2,75)	0,037907	0,036112	0,044296	0,022408	0,024610	0,041280
<2,75; 3,00)	0,063318	0,052051	0,052778	0,045941	0,043341	0,054165
<3,00; 3,25)	0,073398	0,069692	0,060702	0,076109	0,068095	0,067284
<3,25; 3,50)	0,097465	0,086681	0,067391	0,110997	0,095448	0,079124
<3,50; 3,75)	0,121196	0,100148	0,072221	0,124357	0,119359	0,088088
<3,75; 4,00)	0,102042	0,107484	0,074710	0,139018	0,133161	0,092839
<4,00; 4,25)	0,102419	0,107159	0,074601	0,126058	0,132538	0,092630
<4,25; 4,50)	0,093798	0,099241	0,071907	0,107751	0,117691	0,087495
<4,50; 4,75)	0,074792	0,085377	0,066904	0,085236	0,093235	0,078239
<4,75; 5,00)	0,062212	0,068230	0,060088	0,059421	0,065896	0,066232
<5,00; 5,25)	0,040503	0,050651	0,052093	0,038703	0,041549	0,053079
<5,25; 5,50)	0,032387	0,034929	0,043594	0,024514	0,023373	0,040270
<5,50; 5,75)	0,024813	0,022375	0,035215	0,013139	0,011729	0,028923
<5,75; 6,00)	0,013530	0,013314	0,027459	0,007679	0,005251	0,019666
<6,00; 6,25)	0,009445	0,007359	0,020668	0,003776	0,002097	0,012659
<6,25; 6,50)	0,006109	0,003779	0,015017	0,001875	0,000747	0,007714
<6,50; 6,75)	0,003479	0,001802	0,010532	0,000887	0,000238	0,004450
<6,75; 7,00)	0,002140	0,000799	0,007130	0,000377	6,74E-05	0,002430
<7,00; 7,25)	0,001051	0,000329	0,004659	0,000159	1,7E-05	0,001257
<7,25; 7,50)	0,000650	1,26E-04	0,002939	6,56E-05	3,85E-06	0,000615
<7,50; +∞)	0,000724	4,46E-05	0,001790	3,83E-05	7,74E-07	0,000285

Źródło: Badania własne

prawdopodobieństwo, że zmienna losowa przyjmuje wartości z tego przedziału. Dla rozkładu dyskretnego, a takim jest bootstrapowy (zarówno klasyczny jak i dokładny) rozkład estymatora wariancji zmiennej X^D, tak nie jest. Im mniejsza szerokość, tym większej liczbie przedziałów będzie odpowiadało prawdopodobieństwo równe 0.

W tabeli 2 podano również wartości oczekiwane estymatorów wariancji zmiennej X^D oraz ich odchylenia standardowe. Są to wartości dokładne – obliczone. Podane w tabeli rozkłady są tablicowane dla przedziałów. Obliczenie na ich podstawie parametrów rozkładów prowadzić należy jak dla danych pogrupowanych, wobec czego wyniki różnić się mogą od wartości dokładnych.

Wartości oczekiwane obydwu rozkładów granicznych GV1 i GV2 są z założenia równe wariancji próby. W przypadku rozkładu DBV również otrzymano wartość oczekiwaną równą wariancji próby, co potwierdza poprawność użytego algorytmu. Dokładna metoda bootstrapowa nie wprowadza dodatkowego obciążenia estymatora (w przeciwieństwie do metody bootstrapowej z losowaniem prób, gdzie obciążenie jest możliwe).

Na uwagę zasługuje fakt, że odchylenie standardowe rozkładu DBV jest równe z dokładnością do piątego miejsca po przecinku odchyleniu standardowemu rozkładu GV1. Potwierdza to poprawność algorytmu realizującego dokładny bootstrap. W przypadku rozkładu granicznego GV2 odchylenie jest zawyżone. Zawyżenie to wynika z niespełnienia warunku normalności. Różnica jest wyraźna i wskazuje na konieczność obliczania wariancji estymatora ze wzoru (8).

Na wykresie 1 przedstawiono rozkłady estymatorów wariancji zmiennej X^D dla n=20 oraz n=30. Wynika z nich, że rozkład GV2 (rozkład graniczny wymagający założenia normalności) nie stanowi właściwego przybliżenia rozkładu DBV dla rozpatrywanej zmiennej losowej. Oznacza to, że w przypadku estymatorów wariancji niespełnienie warunku normalności nie może być ignorowane. Nie można wówczas stosować rozkładu granicznego, wymagającego założenia, że próba pierwotna ma rozkład normalny. Rozkład GV1 natomiast jest przybliżeniem dobrym.

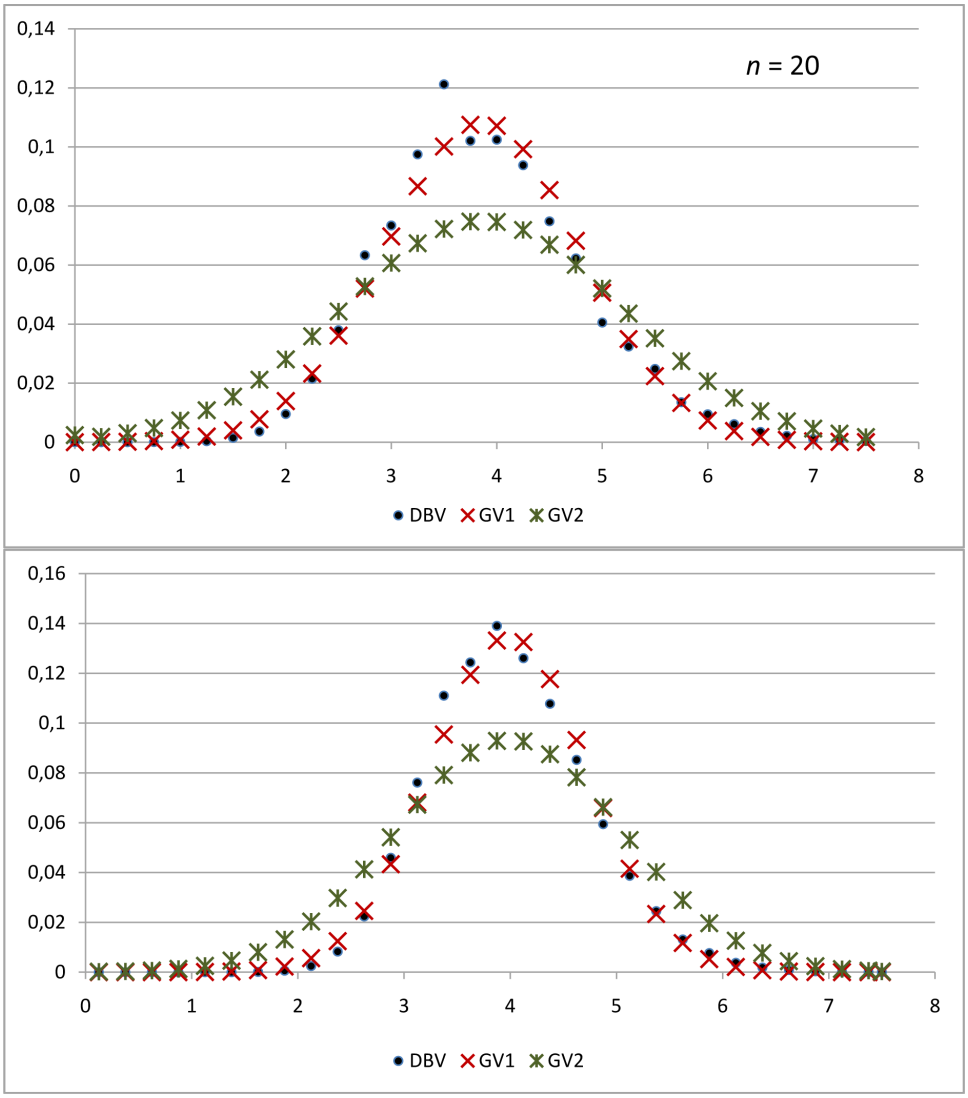
Zauważmy ponadto, że rozkład DBV jest minimalnie asymetryczny, podczas gdy rozkład graniczny jest oczywiście symetryczny.

Zgodność rozkładów GV1 i DBV potwierdził test zgodności Pearsona, zarówno dla n=30 jak dla n=20.

W tabeli 3 przedstawiono przedziały ufności dla wariancji wyznaczone przy użyciu dokładnej metody bootstrapowej DBV, rozkładu granicznego nie wymagającego założenia normalności próby GV1 oraz wymagającego założenia normalności GV2.

Porównując rozstępy poszczególnych przedziałów stwierdzamy, że najprecyzyjniejszego oszacowania dokonano przy pomocy dokładnej metody bootstrapowej (i jest to oszacowanie dokładne), następnie rozkładu granicznego GV1 (z wyjątkiem n=20 i 1-α = 0,99, dla których przedział ufności rozkładu GV1 był węższy od DBV). Najszersze przedziały ufności ma rozkład GV2.

Wykres 1. Rozkłady estymatorów wariancji DBV, GV1 i GV2



Źródło: Badania własne

Przesunięcie w lewo przedziałów dla rozkładu GV1 wobec DBV wynika z asymetrii drugiego z nich. Przedstawione obliczenia wskazują, że rozkład GV1 jest dobrym przybliżeniem rozkładu DBV.

W celu dokładniejszej prezentacji metody przedstawione zostaną wyniki estymacji dla małej próby obejmującej wartości {1, 2, 3, 4, 5} gdzie $n = k$=5 przy założeniu jednakowych prawdopodobieństw wylosowania każdego elementu, równych 0,2. Rozkład ten reprezentuje dyskretna zmienna losowa X^D o wartości oczekiwanej i wariancji

Tabela 3

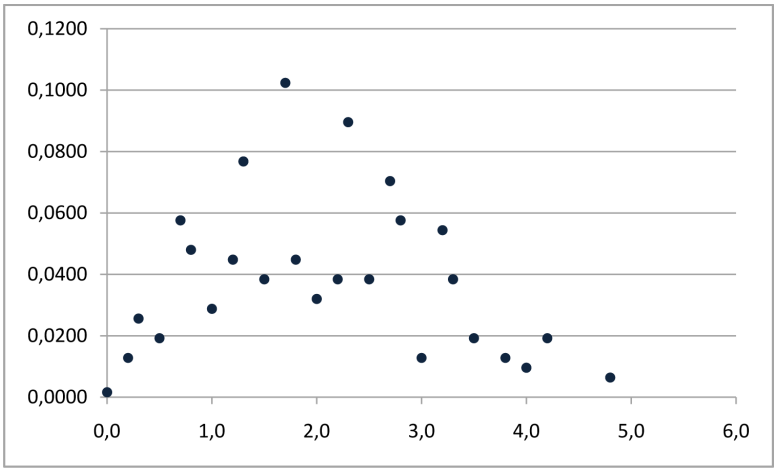
Przedziały ufności dla wariancji wyznaczone dokładną metodą bootstrapową oraz przy użyciu rozkładów granicznych

Poziom ufności		*n*=20			*n*=30		
		DBV	GV1	GV2	DBV	GV1	GV2
1-α = 0,95	**granice**	2,3685	2,1907	1,3868	2,6715	2,5462	1,9011
		5,9565	5,7888	6,5927	5,0845	5,4332	6,0784
	rozstęp	3,5880	3,5981	5,2059	2,4130	2,8870	4,1773
1-α = 0,99	**granice**	1,9575	1,6254	0,5689	2,3265	2,0926	1,2448
		6,6945	6,3541	7,4106	6,1185	5,8868	6,7347
	rozstęp	4,7370	4,7287	6,8417	3,7920	3,7942	5,4899

Źródło: Badania własne

odpowiednio równych μ^D= 3, oraz $\left(\sigma^{\mathrm{D}}\right)^2$= 2. Wartość oczekiwana nieobciążonego estymatora wariancji jest równa 2, a jego wariancja 0,96 (zgodnie ze wzorem (8) po korekcie odpowiednio o czynniki 5/4 i $(5/4)^2$). Z 5 elementowej próby pierwotnej można wylosować 5^5=3125 prób wtórnych, przy czym prawdopodobieństwo wylosowanie każdej z nich jest w zadanych warunkach jednakowe i wynosi 1/3125. Estymator wariancji dla przyjętej próby pierwotnej przybiera jedynie 26 różnych wartości. W tabeli 4 oraz na wykresie 2 przedstawiony został rozkład nieobciążonego estymatora wariancji wyznaczonego dokładną metodą bootstrapową. Wartość oczekiwana i wariancja tego rozkładu są dokładnie równe 2 oraz 0,96, czego należało oczekiwać. Warto dodać, że dla tak małej próby łatwo wykonać niezbędne kalkulacje z poziomu arkusza kalkulacyjnego programu Excel, bez konieczności używania specjalnego oprogramowania.

Wykres 2. Rozkłady estymatora wariancji dla próby obejmującej wartości {1, 2, 3, 4, 5}



Źródło: Badania własne

Tabela 4

Rozkład estymatora wariancji uzyskany dokładną metodą bootstrapową dla 5-cio elementowej próby obejmującej wartości {1, 2, 3, 4, 5}

Wartości	Prawdopodobieństwa	Wartości	Prawdopodobieństwa
0,0	0,0016	2,2	0,0384
0,2	0,0128	2,3	0,0896
0,3	0,0256	2,5	0,0384
0,5	0,0192	2,7	0,0704
0,7	0,0576	2,8	0,0576
0,8	0,0480	3,0	0,0128
1,0	0,0288	3,2	0,0544
1,2	0,0448	3,3	0,0384
1,3	0,0768	3,5	0,0192
1,5	0,0384	3,8	0,0128
1,7	0,1024	4,0	0,0096
1,8	0,0448	4,2	0,0192
2,0	0,0320	4,8	0,0064

Źródło: Badania własne

5. PODSUMOWANIE

W artykule omówiono dokładną metodę bootstrapową, którą można wykorzystać do szacowania estymatorów parametrów zmiennych losowych o nieznanym rozkładzie. Metoda pozwala wyznaczyć oszacowanie dowolnego parametru, błąd tego oszacowania, rozkład estymatora, czy przedziały ufności. Tradycyjnie zadanie takie realizowane jest przy pomocy metody bootstrapowej, która polega na losowaniu prób wtórnych z pierwotnej próby losowej. Losowanie próby stosowane jest w statystyce, jeśli nie może być zbadana cała populacja, lub badanie całej populacji jest zbyt kłopotliwe. Próba pierwotna jest po pierwsze skończona, a po drugie znany jest jej rozkład – jest to rozkład empiryczny. Zamiast wtórnie próbkować próbę pierwotną można wygenerować automatycznie całą przestrzeń prób wtórnych i wyznaczyć dla niej wartości statystyki będącej estymatorem poszukiwanego parametru.

W artykule przedstawiono propozycję algorytmu realizującego to zadanie. Poprawność metody sprawdzono na przykładzie wariancji. Pokazano, że wartość oczekiwana estymatora obliczonego dokładną metodą bootstapową jest równa dokładnie wariancji próby. Metoda nie wprowadza więc obciążenia wynikającego z wtórnego próbkowania jak ma to miejsce w klasycznym bootstrapie. Wniosek ten jest jedynie potwierdzeniem

oczywistego faktu, że badając całą populacje uzyskujemy wyniki dokładniejsze niż jeśli bierzemy pod uwagę próbę losową.

Rozkład nieobciążonego estymatora wariancji wyznaczony dokładną metodą bootstrapową porównano z rozkładem granicznym dla wariancji, nie wymagającym założenia normalności próby pierwotnej. Podobieństwo obydwu rozkładów wskazuje, na możliwość przybliżenia rozkładu dokładnego rozkładem granicznym.

Badania pokazały, że w przypadku estymatora obciążonego nie można pominąć kwestii normalności próby pierwotnej. Rozkład graniczny estymatora wariancji, jeśli próba nie ma rozkładu normalnego, nie może być przybliżony rozkładem $N\left(\sigma^2, \frac{n-1}{n}\sqrt{\frac{2\cdot\sigma^4}{n}}\right)$. Rozkładem granicznym jest rozkład

$$\mathrm{N}\left(\sigma^2, \frac{n}{n-1}\sqrt{\frac{\mu_4-\sigma^4}{n}-\frac{2\cdot(\mu_4-2\cdot\sigma^4)}{n^2}+\frac{\mu_4-3\cdot\sigma^4}{n^3}}\right).$$

Przeprowadzone eksperymenty symulacyjne polegające na estymacji parametrów dla różnych rozmiarów próby pierwotnej pokazały, że w przypadku prób małych ($n \leqslant 15$ i $k = n$) czas niezbędny do wygenerowania całej przestrzeni prób wtórnych jest krótki (poniżej 10 sek. na komputerze średniej klasy). Oznacza to, że nie ma wówczas potrzeby przeprowadzać wtórnego losowania próby. Dla prób większych, czas ten jest zdecydowanie dłuższy i wymaga kilku godzin obliczeń (dla n=20 i $k = n$ obliczenia trwały 5 godz. 30 min.) Wzrost wymiarowości problemu powoduje bardzo silne wydłużenie czasu obliczeń. Z drugiej strony pamiętać należy, że metoda bootstrapowa stosowana jest w przypadku małych prób – dla prób dużych można wykorzystać graniczne rozkłady estymatorów.

Biorąc jednak pod uwagę postęp w technice komputerowej, dokładny bootstrap będzie mógł być w przyszłości stosowany również dla prób większych.

Szkoła Główna Gospodarstwa Wiejskiego

LITERATURA

[1] Boot J.G., Sarkar S. [1998]: *Monte Carlo Approximation of Bootstrap Variances*. The American Statistician. Nr. 52, Assue 4. 354-357.

[2] Boot J.G.: Monte Carlo and the boothstrap. Prezentacja elektroniczna: http://www.stat.ufl.edu/~jbooth/documents/talks/bootse.pdf (2010.1.3).

[3] Domański C., Pruska K. [2000]: *Nieklasyczne metody statystyczne*. Polskie Wydawnictwo Ekonomiczne, Warszawa.

[4] Efron B. [1979]: Bootstrap methods: another look at the jackknife. *The Annals of Statistics*. Vol. 7, No. 1, 1-26.

[5] Efron B. [1987]: Better Bootstrap Confidence Intervals (with discussion). *Journal of the American Statistical Association*. Vol. 82, No. 397, 171-185.

[6] Efron B., Tibshirani R.J. [1993]: *An introduction to the Bootstrap.* Chapman & Hall. London.

[7] Feller W. [1950]: *An introduction to probability theory and its application*. John Wiley & Sons. New York, London, Sydney.
[8] Smirnow N.W., Dunin-Barkowski I.W. [1973]: *Kurs rachunku prawdopodobieństwa i statystyki matematycznej dla zastosowań technicznych*. Państwowe Wydawnictwo Naukowe, Warszawa.
[9] Zieliński R. [2004]: *Siedem wykładów wprowadzających do statystyki matematycznej*. Publikacja elektroniczna: http://www.impan.pl/~rziel/7ALL.pdf (2009.08.09)

DOKŁADNA METODA BOOTSTRAPOWA I JEJ ZASTOSOWANIE DO ESTYMACJI WARIANCJI

Streszczenie

W artykule przedstawiono dokładną metodę bootstrapową, którą można wykorzystać do szacowania estymatorów parametrów zmiennych losowych o nieznanym rozkładzie. Metoda pozwala wyznaczyć oszacowanie dowolnego parametru, błąd tego oszacowania, rozkład estymatora, czy przedziały ufności. Tradycyjnie zadanie takie realizowane jest przy pomocy metody bootstrapowej, która polega na wtórnym próbkowaniu analizowanej, pierwotnej próby losowej. Losowanie próby stosowane jest w statystyce, jeśli nie może być zbadana cała populacja, lub badanie całej populacji jest zbyt kłopotliwe. Próba pierwotna jest po pierwsze skończona, a po drugie znany jest jej rozkład – jest to rozkład empiryczny. Zamiast wtórnie próbkować próbę pierwotną można wygenerować automatycznie całą przestrzeń prób wtórnych i wyznaczyć dla niej wartości statystyki będącej estymatorem poszukiwanego parametru. W artykule przedstawiono propozycję algorytmu realizującego metodę dokładnego bootstrapu, którego poprawność sprawdzono na przykładzie nieobciążonego estymatora wariancji. Pokazano, że wartość oczekiwana estymatora obliczonego dokładną metodą bootstapową jest równa dokładnie wariancji próby. Metoda nie wprowadza więc obciążenia wynikającego z wtórnego próbkowania jak ma to może mieć miejsce w klasycznym bootstrapie. Rozkład estymatora wyznaczony dokładną metodą bootstrapową porównano z rozkładem granicznym estymatora wariancji, nie wymagającym założenia normalności próby pierwotnej. Badania pokazały, że w przypadku wariancji nie można pominąć kwestii normalności próby pierwotnej.

EXACT BOOTSTRAP METHOD AND IT'S APPLICATION IN ESTIMATION OF VARIANCE

Summary

The article presents the exact bootstrap method, which can be used to estimate the parameters of the estimators of random variables with unknown distribution. The method allows to determine an estimate of any parameter, the error of estimation, the distribution of the estimator and confidence intervals. Traditionally this task is carried out using the bootstrap method, which consists of resampling of the original sample. Random sampling is necessary if examining the entire population data is impossible or too costly. Note that the fundamental sample property is of finite size and we know its distribution – it is the empirical distribution. Rather than driving a resample, we can generate automatically the entire resample space and calculate the values of a statistic which is looking for a parameter estimator. This article describes a method for performing exact algorithm for bootstrapping, which correctness was verified on an example of the unbiased estimator of variance. It is shown that the expected value of the

estimator calculated with exact bootstrap is exactly equal the variance of the sample. The method does not introduce bias of the resampling, as it may be for the classic bootstrap. The distribution of the estimator determined by the exact bootstrap compared with the limit distribution for estimator of variance, which does not require assumptions of normality of the original sample. Research has shown that if we estimate variance we cannot ignore the issue of normality of the original sample.

PRZEGLĄD STATYSTYCZNY
R. LVIII – ZESZYT 1-2 – 2011

EMIL PANEK

O PEWNEJ PROSTEJ WERSJI „SŁABEGO" TWIERDZENIA O MAGISTRALI W MODELU VON NEUMANNA

1. WSTĘP

Model J. von Neumanna po raz pierwszy został opublikowany w języku niemieckim w 1928 r., ale szersze zainteresowanie ekonomistów i matematyków wzbudziła dopiero jego angielska wersja [36] z 1945 r. Intensywny okres zainteresowania równowagą v. Neumanna i tzw. efektem magistrali w wieloproduktowych (wielosektorowych) modelach wzrostu typu Neumanna-Gale'a-Leontiefa przypada na drugą połowę XX wieku, zob. np. prace [1, 3, 4-8, 10-35], ale również w pierwszej dekadzie obecnego wieku raz po raz pojawiają się nowe publikacje na tan temat w czołowych czasopismach naukowych, por. [2, 9, 37-39].

W artykule prezentujemy wersję modelu von Neumanna, którego postać nawiązuje do klasycznych prac, [4-7, 12, 16, 24}. Przytaczamy stosunkowo prosty dowód istnienia stanu równowagi neumannowskiej oraz „słabą" wersję twierdzenia o magistrali w takim modelu. Idea dowodu nawiązuje do artykułu R. Radnera [33] (zob. także m.in. H. Nikaido [24] twierdzenie 13.8, A. Takayama [34] twierdzenie 7.A.2).

Zanim przejdziemy do przedstawienia modelu, słów kilka o niektórych symbolach matematycznych stosowanych w pracy. Przez R^n oznaczamy n-wymiarową przestrzeń wektorową (nad ciałem liczb rzeczywistych). Jeżeli wektor $x \in R^n$ jest nieujemny, piszemy $x \geqq 0$. Zapis $x \geqslant 0$ oznacza, że nieujemny wektor x zawiera co najmniej 1 element dodatni, tzn. $x \geqslant 0$ wtedy i tylko wtedy, gdy $x \geqq 0$ oraz $x \neq 0$. Nazywamy go wektorem półdodatnim. Natomiast zapis $x > 0$ oznacza, że wszystkie składowe wektora x są dodatnie. Zapis $x \geqq y$ znaczy tyle, co $x - y \geqq 0$, podobnie $x \geqslant y$ oznacza, że $x - y \geqslant 0$, a $x > y$ oznacza, że $x - y > 0$. Przez R^n_+ oznaczamy nieujemny orthant R^n, tzn. $R^n_+ = \{x \in R^n \,|\, x \geqq 0\}$. Symbolem $\langle x, y\rangle$ oznaczamy iloczyn skalarny wektorów $x, y \in R^n$, czyli $\langle x, y\rangle = \sum_{i=1}^{n} x_i, y_i$. Przez $\|x\|$ oznaczamy tzw. normę taksówkową wektora $x \in R^n$: $\|x\| = \sum_{i=1}^{n} |x_i|$. Jeżeli a jest skalarem, to $a \geqslant 0$ oznacza, że $a = 0$ lub $a > 0$.

2. MODEL VON NEUMANNA. STATYKA

W gospodarce mamy n towarów oraz m procesów produkcji, które nazywamy bazowymi procesami technologicznymi. Skalar $a_{ji} \geqslant 0$ wskazuje na wielkość zużycia i-tego towaru w j-tym bazowym procesie technologicznym prowadzonym z jednostkową intensywnością. Natomiast skalar $b_{ji} \geqslant 0$ wskazuje na produkcję i-tego towaru w j-tym bazowym procesie technologicznym prowadzonym z jednostkową intensywnością. Nieujemne (prostokątne, $m \times n$) macierze

$$A = \begin{bmatrix} a_{11} & ... & a_{1n} \\ a_{21} & ... & a_{2n} \\ ... & ... & ... \\ a_{m1} & ... & a_{mn} \end{bmatrix} \quad , \quad B = \begin{bmatrix} b_{11} & ... & b_{1n} \\ b_{21} & ... & b_{2n} \\ ... & ... & ... \\ b_{m1} & ... & b_{mn} \end{bmatrix}$$

nazywamy (odpowiednio) macierzą nakładów (zużycia) i macierzą wyników (produkcji) w modelu von Neumanna. Zatem w $j-$tym bazowym procesie technologicznym stosowanym z jednostkową intensywnością z wektora (wierszowego) nakładów $a_j = (a_{j1}, a_{j2}, ..., a_{jn})$ można wytworzyć wektor produkcji $b_j = (b_{j1}, b_{j2}, ..., b_{jn})$. Zakładamy, że:

1. **Każdy wiersz nieujemny macierzy A zawiera element dodatni** (tzn. w każdym bazowym procesie technologicznym zużywany jest co najmniej jeden towar).
2. **Każda kolumna nieujemnej macierzy B zawiera element dodatni** (tzn. każdy towar jest wytwarzany w przynajmniej jednym bazowym procesie technologicznym).

Model jest liniowy w tym sensie, że dla dowolnych liczb $\nu_1 \geqslant 0$, $\nu_2 \geqslant 0$, $\ldots \nu_m \geqslant 0$ z nakładów $x = \sum_{j=1}^{m} \nu_j a_j$ otrzymujemy produkcję $y = \sum_{j=1}^{m} \nu_j b_j$.

O parze

$$(x, y) = \left(\sum_{j=1}^{m} \nu_j a_j, \ \sum_{j=1}^{m} \nu_j b_j \right) = (\nu A, \nu B)$$

mówimy, ze opisuje dopuszczalny proces produkcji w modelu von Neumanna. Półdodatni wektor (wierszowy) $\nu = (\nu_1, ..., \nu_m) \geqslant 0$ nazywamy wektorem intensywności stosowania bazowych procesów technologicznych.

Weźmy (niezerowy) wektor intensywności $\nu \geqslant 0$ i utwórzmy proces $(x, y) = (\nu A, \nu B)$. Liczbę

$$\alpha_i(\nu) = \begin{cases} \dfrac{(\nu B)_i}{(\nu A)_i}, & gdy\ (\nu A)_i > 0 \\ +\infty\ , & gdy \quad (\nu A)_i = 0,\ (\nu A)_i > 0 \\ wielkosc\ nieokreslona, & gdy \quad (\nu A)_i = (\nu B)_i = 0 \end{cases}$$

nazywamy wskaźnikiem technologicznej efektywności wytwarzania i-tego towaru w procesie (vA, vB). Liczbę

$$\alpha(v) = \min_i \alpha_i(v)$$

nazywamy wskaźnikiem technologicznej efektywności procesu (vA, vB), a liczbę

$$\alpha_N = \max_{v \geqslant 0} \alpha(v) \tag{1}$$

nazywamy optymalnym wskaźnikiem technologicznej efektywności produkcji w modelu (gospodarce) von Neumanna.

Ponieważ $v \geqslant 0$, więc wprost z definicji mamy

$\alpha(v) = \min_i \alpha_i(v) = \max\{\alpha \,|\alpha\, vA \leqq vB\}$,

czyli zadanie (1) jest równoważne z zadaniem

$$\begin{gathered} \max \alpha \\ \alpha v A \leqq v B \\ v \geqslant 0. \end{gathered} \tag{2}$$

Pokażemy, że przy założeniach (I), (II) zadanie to ma rozwiązanie.[1]

□ **Twierdzenie 1.** Przy założeniach (I), (II) istnieje taki wektor intensywności $\bar{v} \geqslant 0$, że

$$\alpha(\bar{v}) = \max_{v \geqslant 0} \alpha(v) = \alpha_N > 0.$$

Dowód. Pokażemy najpierw, że funkcja $\alpha(v)$ jest wszędzie na R^n_+ poza 0 ciągła oraz dodatnio jednorodna stopnia 0.

Zauważmy, że z definicji

$$\alpha(v) = \min_i \alpha_i(v) = \min_i \frac{\langle v, b^i \rangle}{\langle v, a^i \rangle} = \frac{\langle v, b^k \rangle}{\langle v, a^k \rangle}$$

dla pewnego k.[2]

Funkcja $\alpha(v)$ jest zatem ciągła w każdym takim punkcie $v \geqslant 0$, w którym $\langle v, a^k \rangle > 0$ (jako iloraz funkcji liniowych, a więc ciągłych) . Załóżmy teraz, że

$\alpha(v) = \dfrac{\langle v, b^k \rangle}{\langle v, a^k \rangle}$ oraz $\langle v, a^k \rangle = 0$.

---

[1] Inny dowód przytacza m.in. D. Gale [7], twierdzenie 9.8.

[2] Symbolem a^k, b^k oznaczamy k-tą kolumnę macierzy Ai B:

$$a^k = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix}, \; b^k = \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{mk} \end{pmatrix}$$

Przy założeniu (I) warunek $v \neq 0$ pociąga za sobą $vA \geqslant 0$, czyli $\langle v, a^j \rangle > 0$ dla pewnego j, tzn.

$$\alpha(v) = \min_i \frac{\langle v, b^i \rangle}{\langle v, a^i \rangle} = \frac{\langle v, b^k \rangle}{\langle v, a^k \rangle} \leqslant \alpha_j(v) < +\infty \quad ,$$

co jest niemożliwe. Jeżeli bowiem $\langle v, b^k \rangle > 0$, to $\alpha(v) = +\infty$, jeżeli zaś $\langle v, b^k \rangle = 0$, to w punkcie v funkcja $\alpha(v)$ jest nieokreślona . Zatem, gdy $v \geqslant 0$, to

$\alpha(v) = \min_i \frac{\langle v, b^i \rangle}{\langle v, a^i \rangle} = \frac{\langle v, b^k \rangle}{\langle v, a^k \rangle}$ dla takiego (pewnego) k że $\langle v, a^k \rangle > 0$

i tym samym funkcja $\alpha(v)$ jest ciągła wszędzie na R^n_+ poza 0.

Ponieważ dla każdego $v \geqslant 0$ istnieje takie $k \in \{1, ..., n\}$, że $\langle v, a^k \rangle > 0$ i $\alpha(v) = \frac{\langle v, b^k \rangle}{\langle v, a^k \rangle}$, więc dla dowolnej liczby $\lambda > 0$ mamy

$$\alpha(\lambda v) = \frac{\langle \lambda v, b^k \rangle}{\langle \lambda v, a^k \rangle} = \frac{\langle v, b^k \rangle}{\langle v, a^k \rangle} = \alpha(v),$$

co dowodzi dodatniej jednorodności stopnia 0 funkcji $\alpha(v)$ wszędzie na R^n_+ poza 0.

Wówczas

$$\max_{v \geqslant 0} \alpha(v) = \max_{\substack{v \geqslant 0 \\ \|v\| = 1}} \alpha(v). \tag{3}$$

Zbiór $S^n_+(1) = \{v \geqslant 0 \,|\, \|v\| = 1\}$ jest zwarty oraz funkcja $\alpha(v)$ jest ciągła na $S^n_+(1)$, więc zadanie (3) ma rozwiązanie $\bar{v} \geqslant 0$. Macierze A, B spełniają warunki (I), (II), więc istnieje taka liczba $\sigma > 0$, że

$\sigma eA \leqq eB > 0$,

gdzie $e = (1, ..., 1)$ i wobec tego $\alpha_N = \alpha(\bar{v}) \geqslant \sigma > 0$.

Wektor $\bar{v} \geqslant 0$, dla którego $\alpha_N = \alpha(\bar{v})$, nazywamy optymalnym wektorem intensywności w modelu von Neumanna. Wektory $\bar{x} = \bar{v}A$, $\bar{y} = \bar{v}B$ nazywamy optymalnymi wektorami nakładów (zużycia) i wyników (produkcji). O parze $(\bar{x}, \bar{y}) = (\bar{v}A, \bar{v}B)$ mówimy, że opisuje optymalny proces produkcji. Przy założeniach (I), (II), zgodnie z twierdzeniem 1, optymalne procesy produkcji w modelu von Neumanna istnieją i są określone z dokładnością do mnożenia przez dowolną stałą dodatnią (z dokładnością do struktury).

Trzecie założenie brzmi następująco:

(III) Istnieje taki optymalny wektor intensywności $\bar{v} \geqslant 0$, **któremu odpowiada wektor produkcji**

$$\bar{y} = \bar{v}B > 0.$$

Innymi słowy, wśród optymalnych procesów produkcji istnieje przynajmniej jeden proces, w którym wytwarzane są wszystkie towary. Model von Neumanna spełniający

ten warunek nazywamy regularnym. Jeżeli spełnione jest założenie (III), to spełnione jest także założenie (II), dlatego w przytoczonych dalej twierdzeniach 2 i 3 jest ono pomijane.

Niech $p \geqslant 0$ oznacza n-wymiarowy (kolumnowy) wektor cen towarów. Liczbę

$$\beta(v,p) = \begin{cases} \dfrac{vBp}{vAp}, & gdy \quad vAp > 0 \\ +\infty, & gdy\ vAp = 0, \quad oraz \quad vBp > 0 \\ wielkosc\ nieokreslona, & gdy \quad vAp = vBp = 0 \end{cases}$$

nazywamy wskaźnikiem ekonomicznej efektywności procesu (vA, vB) przy cenach p.

Δ **Definicja 1.** Mówimy, że gospodarka z technologiczną efektywnością $\alpha(\bar{v}) = \alpha_N > 0$ znajduje się w równowadze von Neumanna, jeżeli obowiązują w niej takie ceny $\bar{p} \geqslant 0$, przy których dla dowolnego wektora intensywności $v \geqslant 0$ spełniony jest warunek:

$$vB\bar{p} - \alpha_N vA\bar{p} \leqq 0. \tag{4}$$

▲

Wektor $\bar{p}$ nazywamy wektorem cen von Neumanna. Łatwo zauważyć, że ceny von Neumanna są określone z dokładnością do struktury. O trójce $\{\alpha_N, \bar{v}, \bar{p}\}$, w której wektor intensywności $\bar{v} \geqslant 0$ spełnia założenie (III), mówimy, że tworzy (optymalny) stan równowagi neumannowskiej.

Zgodnie z twierdzeniem 1:

$$\alpha_N \bar{v}A \leqq \bar{v}B. \tag{5}$$

Z (4) otrzymujemy (dla $v = \bar{v}$)

$$\bar{v}B\bar{p} \leqq \alpha_N \bar{v}A\bar{p},$$

a z (5) (po przemnożeniu obu stron nierówności przez wektor $\bar{p} \geqslant 0$):

$$\bar{v}B\bar{p} \geqq \alpha_N \bar{v}A\bar{p},$$

tzn.

$$\alpha_N \bar{v}A\,\bar{p} = \quad \bar{v}\,B\,\bar{p} \quad ,$$

przy czym $\bar{v}\,B\,\bar{p} > 0$, czyli także $\bar{v}A\,\bar{p} > 0$. Tak więc wszędzie gdzie funkcja $\beta(\cdot, \bar{p})$ jest określona mamy

$$\beta(v,\bar{p}) = \frac{vB\bar{p}}{vA\bar{p}} \leqslant \alpha(\bar{v}) = \alpha_N$$

oraz

$$\beta(\bar{v},\bar{p}) = \frac{\bar{v}B\bar{p}}{\bar{v}A\bar{p}} = \max_{v \geqslant 0} \frac{vB\bar{p}}{vA\bar{p}} = \alpha(\bar{v}) = \alpha_N \quad .$$

W równowadze von Neumanna ekonomiczna efektywność produkcji jest równa efektywności technologicznej i jest to maksymalna efektywność, jaką może osiągnąć gospodarka.

□ **Twierdzenie 2.** Jeżeli spełnione są założenia (I), (III), to istnieją ceny von Neumanna

Dowód. Przy założeniach (I), (III) zbiór

$Q = \{q \mid q = v(\alpha_N A - B),\ v \geqq 0\}$

jest nietrywialnym stożkiem wielościennym w R^n z wierzchołkiem w 0, który nie zawiera wektorów ujemnych. Rzeczywiście, załóżmy że $q \in Q$ i $q < 0$. Wówczas $q = v(\alpha_N A - B) = \alpha_N v A - v B < 0$ dla pewnego wektora $v \geqslant 0$, co jest niemożliwe (sprzeczne z definicją optymalnego wskaźnika technologicznej efektywności α_N).

Utwórzmy zbiór

$$D = Q + R^n_+ = \{d \mid d = q + r,\ q \in Q,\ r \in R^n_+\} \quad .$$

Wówczas $D \neq R^n$ (gdyż zbiór D nie zawiera wektorów ujemnych) oraz $e_i \in D$, $i = 1, ..., n$. Zatem istnieje taka hiperpłaszczyzna z wektorem kierunkowym $\bar{p} \neq 0$, że

$$\langle \bar{p}, d \rangle \geqslant 0 \text{ dla każdego } d \in D. \tag{6}$$

Do zbioru D należą w szczególności wektory $e_i = (0, ..., \underset{(i)}{1}, ...0)$, $i = 1, ..., n$, więc $\langle \bar{p}, e_i \rangle = \bar{p}_i \geqslant 0$ dla każdego i, a ponieważ $\bar{p} \neq 0$, więc $\bar{p} \geqslant 0$. Z (6) wynika w szczególności, że dla każdego $v \geqslant 0$:

$$v(\alpha_N A - B)\bar{p} \geqslant 0 \quad ,$$

co jest równoważne (4).

■

Podobnie jak optymalny wektor intensywności $\bar{v} \geqslant 0$, również ceny von Neumanna $\bar{p} \geqslant 0$ są określone z dokładnością do struktury.

3. MODEL VON NEUMANNA. DYNAMIKA

Załóżmy, że czas t zmienia się skokowo, $t = 0, 1, ..., t_1$. Symbolem $v(t) = (v_1(t), ..., v_n(t))$ oznaczamy wektor (wierszowy) intensywności, z jaką poszczególne bazowe procesy technologiczne są stosowane w okresie t. Model von Neumanna opisuje gospodarkę zamkniętą w tym sensie, że źródłem nakładów w okresie t+1 może być tylko produkcja wytworzona w okresie poprzednim t, co prowadzi do nierówności:

$$v(t+1)A \leqq v(t)B, \quad t = 0, 1, ..., t_1 - 1 \tag{7}$$

$$v(t) \geqslant 0, \quad t = 0, 1, ..., t_1.$$

Niech v^0 będzie dowolnym wektorem intensywności stosowania bazowych procesów technologicznych w okresie $t = 0$

$$v(0) = v^0 \geqslant 0. \tag{8}$$

Δ **Definicja 2.** Ciąg wektorów $\{v(t)\}_{t=0}^{t_1}$ spełniających warunki (7)-(8) nazywamy $(v^0 t_1)$ – dopuszczalną trajektorią intensywności. Ciągi $\{x(t)\}_{t=0}^{t_1}$, $\{y(t)\}_{t=0}^{t_1}$, gdzie $x(t) = v(t)A$, $y(t) = v(t)B$, nazywamy (odpowiednio) dopuszczalną trajektorią nakładów (zużycia) i wyników (produkcji). O trójce $\{v(t), x(t), y(t)\}_{t=0}^{t_1}$ mówimy, że opisuje dopuszczalny proces wzrostu w modelu von Neumanna.

▲

Spośród dopuszczalnych procesów wzrostu szczególną klasę tworzą procesy stacjonarne.

Δ **Definicja 3.** $(v^0 t_1)$ – dopuszczalną trajektorię intensywności $\{v(t)\}_{t=0}^{t_1}$ nazywamy stacjonarną, jeżeli dla pewnej liczby $\gamma > 0$:

$$v(t) = \gamma^t v^0, \quad t = 0, 1, ..., t_1. \tag{9}$$

Stacjonarnej trajektorii intensywności odpowiada stacjonarna trajektoria nakładów

$$x(t) = \gamma^t x^0 \tag{10}$$

oraz wyników

$$y(t) = \gamma^t y^0, \tag{11}$$

gdzie $x^0 = v^0 A$, $y^0 = v^0 B$. O trójce trajektorii (9) – (11) mówimy, że opisują stacjonarny proces wzrostu w modelu von Neumanna.

▲

Warunkiem koniecznym i dostatecznym istnienia stacjonarnego procesu wzrostu z tempem $\gamma > 0$ i początkowym wektorem intensywności $v^0 \geqslant 0$ jest spełnienie układu nierówności:

$$\gamma v^0 A \leqq v^0 B. \tag{12}$$

Przy założeniach (I), (III), zgodnie z twierdzeniem 1, istnieje takie rozwiązanie układu (12) z tempem $\gamma = \alpha_N > 0$ i określonym z dokładnością do struktury wektorem $v^0 = \bar{v} \geqslant 0$, że $\bar{v}B > 0$. Innymi słowy istnieje stacjonarny proces wzrostu $\{\bar{v}(t), \bar{x}(t), \bar{y}(t)\}_{t=0}^{t_1}$ postaci:

$$\bar{v}(t) = \alpha_N^t \bar{v}, \tag{13}$$

$$\bar{x}(t) = \alpha_N^t \bar{x}, \tag{14}$$

$$\bar{y}(t) = \alpha_N^t \bar{y}, \tag{15}$$

z wektorami $\bar{x} = \bar{v}A \geqslant 0$, $\bar{y} = \bar{v}B > 0$. Nazywamy go optymalnym stacjonarnym procesem wzrostu w modelu von Neumanna. Trajektorie (13) – (15) nazywamy, odpowiednio, optymalna stacjonarną trajektorią intensywności, nakładów (zużycia) i wyników (produkcji).

O wektorach

$$\bar{s}^v = \frac{\bar{v}(t)}{\|\bar{v}(t)\|} = \frac{\alpha_N^t \bar{v}}{\left\|\alpha_N^t \bar{v}\right\|} = \frac{\bar{v}}{\|\bar{v}\|} = \left(\frac{\bar{v}_1}{\sum\limits_i \bar{v}_i}, ..., \frac{\bar{v}_m}{\sum\limits_i \bar{v}_i} \right) = const.,$$

$$\bar{s}^x = \frac{\bar{x}(t)}{\|\bar{x}(t)\|} = \frac{\alpha_N^t \bar{x}}{\left\|\alpha_N^t \bar{x}\right\|} = \frac{\bar{x}}{\|\bar{x}\|} = \left(\frac{\bar{x}_1}{\sum\limits_i \bar{x}_i}, ..., \frac{\bar{x}_n}{\sum\limits_i \bar{x}_i}\right) = const.,$$

$$\bar{s}^y = \frac{\bar{y}(t)}{\|\bar{y}(t)\|} = \frac{\alpha_N^t \bar{y}}{\left\|\alpha_N^t \bar{y}\right\|} = \frac{\bar{y}}{\|\bar{y}\|} = \left(\frac{\bar{y}_1}{\sum\limits_i \bar{y}_i}, ..., \frac{\bar{y}_n}{\sum\limits_i \bar{y}_i}\right) = const.,$$

mówimy, że charakteryzują strukturę intensywności stosowania bazowych procesów technologicznych, strukturę nakładów i strukturę produkcji w optymalnym stacjonarnym procesie wzrostu. O półprostych

$$N^v = \{\lambda \bar{s}^v \mid \lambda > 0\},$$

$$N^x = \{\lambda \bar{s}^x \mid \lambda > 0\},$$

$$N^y = \{\lambda \bar{s}^y \mid \lambda > 0\},$$

mówimy, że tworzą magistrale (promienie von Neumanna) w przestrzeni intensywności, przestrzeni nakładów i produkcji. Wszystkie optymalne stacjonarne procesy wzrostu przebiegają po magistralach.

W celu zapewnienia jednoznaczności magistral przyjmiemy następujące założenie:

(IV) Dla każdego wektora intensywności $v \notin N^v$ zachodzi nierówność:

$$v\,B\,\bar{p} - \alpha_N v A \bar{p} < 0.$$

Łatwo zauważyć, że warunek ten zachodzi także dla dowolnego wektora $x = vA \notin N^x$ oraz $y = vB \notin N^y$. Założenie głosi zatem, że jedynie na magistralach ekonomiczna efektywność produkcji jest równa efektywności technologicznej (wszędzie poza magistralami efektywność ekonomiczna jest niższa od technologicznej).[3]

Ustalmy początkowy wektor intensywności $v^0 > 0$ i postawmy następujące (liniowe) zadanie programowania dynamicznego:

$$\begin{gathered} \max v(t_1) B \bar{p} \\ v(t+1)A \leqq v(t)B, \quad t = 0, 1, ..., t_1 - 1 \\ v(0) = v^0 (\text{dane}). \end{gathered} \tag{16}$$

Jest to zadanie maksymalizacji wartości produkcji w końcowym okresie horyzontu $\{0, 1, ..., t_1\}$, mierzonej w cenach von Neumanna $\bar{p} \geqslant 0$, na zbiorze wszystkich (v^0, t_1)-dopuszczalnych trajektorii intensywności. Jego rozwiązanie $\{v^*(t)\}_{t=0}^{t_1}$ nazywamy $(v^0, t_1, \bar{p})$- optymalną trajektorią intensywności. Ciągi $\{x^*(t)\}_{t=0}^{t_1}$, $\{y^*(t)\}_{t=0}^{t_1}$,

[3] Innymi słowy przy założeniu (III) w żadnej przestrzeni – intensywności, nakładów, produkcji – nie ma dwóch różnych magistral (o różnej strukturze). Założenie to znacznie ułatwia dowód twierdzenia 3, choć nie jest generalnie wymagane przy dowodach innych twierdzeń o magistrali formułowanych w literaturze, por. prace wymienione we wstępie.

gdzie $x^*(t) = v^*(t)A$ oraz $y^*(t) = v^*(t)B$, nazywamy $(x^0, t_1, \bar{p})$- optymalną trajektorią nakładów (z początkowym wektorem nakładów $x^0 = v^0A$) oraz $(y^0, t_1, \bar{p})$-optymalną trajektorią produkcji (z początkowym wektorem produkcji $y^0 = v^0B$). O trójce $\{v^*(t), x^*(t), y^*(t)\}_{t=0}^{t_1}$ mówimy, że opisuje optymalny proces wzrostu w modelu von Neumanna.

4. „SŁABE" TWIERDZENIE O MAGISTRALI

W literaturze znanych jest wiele twierdzeń o stabilności typu magistralnego optymalnych procesów wzrostu w modelach von Neumanna-Gale'a-Leontiefa (por. załączoną bibliografię). Choć niekiedy modele te znacznie różnią się między sobą, istota dowodzonych na ich gruncie twierdzeń pozostaje ta sama. Wszystkie głoszą, że bez względu na początkowy stan gospodarki, procesy wzrostu optymalne w sensie obszernej klasy kryteriów wzrostu (nie tylko kryterium maksymalizacji wartości produkcji w końcowym okresie ustalonego horyzontu) „prawie zawsze" przebiegają w dowolnie bliskim otoczeniu magistral. Zbieżność jest tym wyraźniejsza, im dłuższy jest zakładany horyzont gospodarki. Magistrale jawią się jako „drogi szybkiego ruchu" do których powinna zmierzać (ewentualnie, po których powinna poruszać się) każda racjonalnie rozwijająca się gospodarka. Tempo wzrostu na magistralach jest bowiem maksymalnym tempem możliwym do osiągnięcia przez gospodarkę.

□ **Twierdzenie 3 (Radner).** Jeżeli spełnione są założenia (I), (III) i (IV) to dla dowolnej liczby $\varepsilon > 0$ istnieje taka liczba naturalna $k_\varepsilon > 0$, że liczba okresów czasu, w których zachodzi choćby jeden z warunków

$$\left\| \frac{v^*(t)}{\|v^*(t)\|} - \bar{s}^v \right\| \geqslant \varepsilon, \quad \left\| \frac{x^*(t)}{\|x^*(t)\|} - \bar{s}^x \right\| \geqslant \varepsilon, \quad \left\| \frac{y^*(t)}{\|y^*(t)\|} - \bar{s}^s \right\| \geqslant \varepsilon \tag{17}$$

nie przekracza k_ε. Liczba k_ε nie zależy od długości horyzontu t_1.

Dowód. Pokażemy najpierw, że dla dowolnej liczby $\varepsilon > 0$ istnieje taka liczba $\delta_\varepsilon^v > 0$, iż warunek

$$\left\| \frac{v}{\|v\|} - \bar{s}^v \right\| \geqslant \varepsilon \tag{18}$$

pociąga za sobą nierówność

$$vB\bar{p} - (\alpha_N - \delta_\varepsilon^v)vA\bar{p} \leqslant 0. \tag{19}$$

Przy dowodzie wystarczy ograniczyć się do wektorów intensywności $v \geqslant 0$ unormowanych do 1, gdyż jeżeli jakikolwiek wektor $v \geqslant 0$ spełnia warunki (18), (19), to spełnia je także wektor λv, gdzie λ jest dowolną liczbą dodatnią . Zbiór

$$V_\varepsilon = \{v \in S_+^n(1) \mid \|v - \bar{s}^v\| \geqslant \varepsilon\}$$

jest zwarty oraz dla każdego $v \in V_\varepsilon$

$$v \notin N^v,$$

czyli

$$vB\bar{p} - \alpha_N vA\bar{p} < 0 \tag{20}$$

(zgodnie z założeniem (III)), tzn.

$$vA\bar{p} > 0 \quad \text{dla każdego} \quad v \in V_\varepsilon.$$

Ponieważ $\beta(v, \bar{p}) = \dfrac{vB\bar{p}}{vA\bar{p}}$ jest ciągłą na V_ε funkcją zmiennej v, więc (wobec (20)) istnieje liczba

$$\bar{\beta} = \max_{v \in V_\varepsilon} \beta(v, \bar{p}) < \alpha_N,$$

czyli $\alpha_N - \delta_\varepsilon^v \geqslant \bar{\beta}$ dla pewnej liczby $\delta_\varepsilon^v > 0$, skąd wnioskujemy, że dla każdego $v \in V_\varepsilon$

$$\alpha_N - \delta_\varepsilon^v \geqslant \frac{vB\bar{p}}{vA\bar{p}},$$

co prowadzi do warunku (19).

Podobnie dowodzi się, że dla dowolnej liczby $\varepsilon > 0$ istnieją takie liczby $\delta_\varepsilon^x > 0$, $\delta_\varepsilon^y > 0$, iż nierówność $\left\| \dfrac{x}{\|x\|} - \bar{s}^x \right\| \geqslant \varepsilon$ pociąga za sobą $vB\bar{p} - (\alpha_N - \delta_\varepsilon^x) vA\bar{p} \leqslant 0$, a nierówność $\left\| \dfrac{y}{\|y\|} - \bar{s}^y \right\| \geqslant \varepsilon$ pociąga za sobą nierówność $vB\bar{p} - (\alpha_N - \delta_\varepsilon^y) vA\bar{p} \leqslant 0$, gdzie $x = vA$, $y = vB$. Biorąc $\delta_\varepsilon = \min\{\delta_\varepsilon^v, \delta_\varepsilon^x, \delta_\varepsilon^y\}$ dochodzimy do wniosku, iż dla dowolnej liczby $\varepsilon > 0$ istnieje taka liczba $\delta_\varepsilon > 0$, że jeżeli zachodzi którakolwiek z nierówności

$$\left\| \frac{v}{\|v\|} - \bar{s}^v \right\| \geqslant \varepsilon, \quad \left\| \frac{x}{\|x\|} - \bar{s}^x \right\| \geqslant \varepsilon, \quad \left\| \frac{y}{\|y\|} - \bar{s}^y \right\| \geqslant \varepsilon$$

(gdzie $x = vA$, $y = vB$), to

$$vB\bar{p} - (\alpha_N - \delta_\varepsilon^{)}) vA\bar{p} \leqslant 0. \tag{21}$$

Niech $\{v^*(t)\}_{t=0}^{t_1}$ będzie $(v^0, t_1, \bar{p})$- optymalną trajektorią intensywności rozwiązaniem zadania (16). Wówczas, zgodnie z (4)

$$v^*(t)\, B\, \bar{p} \leqslant \alpha_N v^*(t) A\, \bar{p}, \quad t = 0, 1, ..., t_1$$

oraz

$$v^*(t+1)\, A \leqslant v^*(t) B, \quad t = 0, 1, ..., t_1 - 1$$

(w myśl (7)), czyli

$$v^*(t+1)\, A\, \bar{p} \leqslant v^*(t) AB\bar{p} \leqslant \alpha_N v^*(t) A\, \bar{p}, \quad t = 0, 1, ..., t_1 - 1. \tag{22}$$

Niech L będzie zbiorem tych okresów czasu, w których zachodzi którykolwiek z warunków (17). Niech n_ε będzie liczbą elementów zbioru L. Wówczas

$$v^*(t+1)\,A\,\bar{p} \leqslant (\alpha_M - \delta_\varepsilon)v^*(t)A\bar{p} \quad \text{dla } t \in L. \tag{23}$$

Z (22), (23) otrzymujemy nierówność

$$v^*(t_1)A\bar{p} \leqslant \alpha_N^{t_1-n_\varepsilon}(\alpha_N - \delta_\varepsilon)^{n_\varepsilon}v^0 A\,\bar{p}. \tag{24}$$

Ponieważ $v^0 > 0$, więc (przy założeniu (III)) istnieje taka liczba $\sigma > 0$, że
$0 < \sigma\,\bar{s}^v A \leqq v^0 B$,
i wobec tego istnieje $(v^0 t_1)$ – dopuszczalna trajektoria $\{\tilde{v}(t)\,\}_{t=0}^{\tau}$:

$$\tilde{v}(t) = \begin{cases} v^0, & dla \quad t = 0 \\ \alpha_N^{t-1}\sigma\bar{s}^v, & dla \quad t = 1, ..., t_1. \end{cases}$$

Wówczas

$$v^*(t_1)\,B\,\bar{p} \geqslant \tilde{v}(t_1)\,B\,\bar{p} = \alpha_N^{t_1-1}\sigma\bar{s}^v B\bar{p} > 0. \tag{25}$$

Łącząc warunki (24) i (25) (zważywszy, że $v^*(t_1)\,B\,\bar{p} \leqslant \alpha_N v^*(t_1)A\,\bar{p}$) otrzymujemy:
$$0 < \alpha_N^{t_1-1}\sigma\bar{s}^v B\bar{p} \leqslant v^*(t_1)\,B\,\bar{p} \leqslant \alpha_N v^*(t_1)A\,\bar{p} \leqslant \alpha_N^{t_1-n_\varepsilon+1}(\alpha_N - \delta_\varepsilon)^{n_\varepsilon}v^0 A\,\bar{p},$$

co pozwala na oszacowanie górnego ograniczenia liczby n_ε:

$$n_\varepsilon \leqslant \mu = \frac{\ln A}{\ln\alpha - \ln(\alpha_N - \delta_\varepsilon)},$$

gdzie $A = \dfrac{\alpha_N^2 v^0 A\bar{p}}{\sigma\bar{s}^v B\bar{p}} > 0$. W charakterze liczby k_ε wystarczy wziąć najmniejszą liczbę naturalną większą od $\min\{0, \mu\}$.

■

LITERATURA

[1] Araujo A., Scheinkman J.A. [1977], *Smoothness, Comparative Dynamics and Turnpike Property*, „Econometrica", 43/1977.

[2] Blot J., Crettez B., [2007], *On the Smoothness of Optimal Path II: Some Local Turnpike Results,* „Decision Econom. Finance", nr 2/2007.

[3] *Contribution to the von Neumann Growth Model*, (Bruckman G., Weber W. – red.) [1971], Springer Verlag, New York -Wien.

[4] Czeremnych J., N. [1982], *Analiz powiedienija trajektorii dynamiki narodnochaziajstwiennych modeliej*, Nauka, Moskwa.

[5] Czeremnych J., N. [1986], *Matiematiczieskije modeli razwitija narodnogo chaziajstwa*, MGU, Moskwa.

[6] Czerwiński Z., [1973], *Podstawy matematycznych modeli wzrostu gospodarczego*, PWE, Warszawa.

[7] Gale D., [1969], *Teoria liniowych modeli ekonomicznych*, PWN, Warszawa.

[8] Gantz D.T. [1980], *A Strong Turnpike Theorem for Nonstationary von Neumann-Gale Production Model*, „Econometrica" t. 48, nr 7/1980.
[9] Guerrero-Luchtenberg C.L. [2000], *A Uniform Neighborhood Turnpike Theorem and Applications*, „Journal of Math. Econ." 34/2000.
[10] *Handbook of Mathematical Economics*, (Arrow K.J., Intrigilator M.D. – red.) [1982], t.II, North-Holland, Amsterdam-New York-Oxford.
[11] *John von Neumann and Modern Economics*, (Dove M., Chakravorty S., Goodwin R. - red.) [1989], Clarendon Press, Oxford.
[12] Kemeny J.G., Morgenstern O., Thompson G.L. [1956], *A Generalization of the von Neumann Model of Expanding Economy,* „Econometrica" 24/1956.
[13] Krass J.A. [1976], *Matiematiczieskije modeli ekonomiczeskoj dinamiki*, Sowietskoje Radio, Moskwa.
[14] Makarow W.L., Rubinow A.M. [1973], *Matiematiczeskaja tieoria ekonomiczeskoj dinamiki i rawnowiesija*, Nauka, Moskwa.
[15] Marena M., Montrucchio L., [1999], *Neighborhood Turnpike Theorem for Continous Time Optimization Models*, Journal Optim. Theory Appl. 101/1999.
[16] *Mathematical Models of Economics*, (Łoś J., Łoś M. – red.) [1974], North-Holland, Amsterdam.
[17] McKenzie L.W. [1976], *Turnpike Theory*, „Econometrica", t.44, nr 5/1976.
[18] McKenzie L.W. [1998], *Turnpikes*, „American Econ. Rev. 88 (22)/1998.
[19] Montrucchio L. [1995], *A New Turnpike Theorem for Discounted Programs*, „Econ. Theory", 5/1995.
[20] Morishima M. [1964], *Equilibrium, Stability and Growth*, Clarendon Press, Oxford.
[21] Morishima M. [1969], *Theory of Economic Growth*, Clavendon Press, Oxford.
[22] Mowszowicz S.M. [1969], *Tieoriemy o magistrali w modeliach Neumanna-Gale'a*, „Ekonomika i matiematiczeskije metody", nr 6/1969.
[23] Mowszowicz S.M. [1972], *Magistralnyj rost w dinamiczeskich narodnochoziastwiennych modeliach*, „Ekonomika i matiematiczeskije metody", nr 2/1972.
[24] Nikaido H., [1968], *Convex Structers and Economic Theory*, Acad. Press, New York.
[25] Panek E., [1985], *Asymptotyka optymalnych trajektorii w wielosektorowym modelu wzrostu*, „Przegląd statystyczny" nr 2/1985.
[26] Panek E., [1985], *Słabe twierdzenie o magistrali konsumpcyjnej w wielosektorowym modelu wzrostu*, Przegląd Statystyczny nr 3/1985.
[27] Panek E., [1986], *Asymptotyka trajektorii produkcji w dynamicznym modelu Leontiefa,* Przegląd Statystyczny nr 2/1986.
[28] Panek E., [1987], *O pewnej wersji twierdzenia o magistrali w wielosektorowym modelu wzrostu,* Przegląd Statystyczny nr 1/1987.
[29] Panek E., [1988], *Consumption Turnpike in a Nonlinear Model of Input-Output Type*, w: Input – Output Analysis. Current Developments, red. M. Chiaschini, Chapman and Hall Ltd., London-New York.
[30] Panek E., [1992], *Optimal Trajectories in a Multisectoral Model of Economic Growth*, Computers and mathematics with Applications, nr 8-9 (24)/1992.
[31] Panek E., [1992], *„Bardzo silne" twierdzenie o magistrali w wielosektorowym modelu wzrostu*, Przegląd Statystyczny nr 3-4/1992.
[32] Panek E., [1997], *"Silne" twierdzenie o magistrali w wielosektorowym modelu wzrostu*, *Wersja szczególna,* Przegląd Statystyczny nr 1/1997.
[33] Radner R.[1961], *Path of Economic Growth that are Optimal with Regard only to Final States: A Turnpike Theorem*, Review of Econ. Studies, XXVIII, 1961.
[34] Takayama A. [1985], *Mathematical Economics*, Cambridge Univ. Press, Cambridge.
[35] Tsukui J., Murokami Y. [1978], *Turnpike Optimality in Input-Output Systems. Theory and Applications for Planning*, North-Holland, Amsterdam-New York-Oxford.

[36] Von Neumann J. [1945], *A Model of General Economic Equilibrium*, Rev. Econ. Studies, 13/1945-46.
[37] Zaslawski A.J. [2000], *Turnpike Theorem for Nonautonomous Infinite Dimensional Discrete-Time Control Systems,* Optimization, 48/2000.
[38] Zaslawski A.J. [2004], *A Turnpike Result for Autonomous Variational Problems*, Optimization, 53/2004.
[39] Zaslawski A.J. [2006], *Turnpike Properties in the Calculus of Variations and Optimal Control*, Springer Science & Business Media, Inc New York.

O PEWNEJ PROSTEJ WERSJI „SŁABEGO" TWIERDZENIA O MAGISTRALI W MODELU VON NEUMANNA

S t r e s z c z e n i e

Prezentujemy klasyczną wersję modelu von Neumanna, którego postać nawiązuje do prac [4-7, 12, 16, 24]. Przytaczamy prosty dowód istnienia stanu równowagi oraz „słabą" wersję twierdzenia R. Radnera o magistrali w takim modelu.

Słowa kluczowe: równowaga von Neumanna, stacjonarny i optymalny proces wzrostu, efekt magistrali

JEL codes: O 41

A SIMPLE VERSION OF "WEAK" TURNPIKE THEOREM IN THE VON NEUMANN MODEL

S u m m a r y

In the paper we present a classical version of von Neumann model which refers to works [4-7, 12, 16, 24]. We adduce a simple proof of equilibrium state and a "weak" version of R. Radner theorem – about turnpike in such model.

Key words: von Neumann theorem, stationary and optimal growth process, turnpike effect (result)

PRZEGLĄD STATYSTYCZNY
R. LVIII – ZESZYT 1-2 – 2011

EMILIA TOMCZYK

APPLICATION OF MEASURES OF ENTROPY, INFORMATION CONTENT AND DISSIMILARITY OF STRUCTURES TO BUSINESS TENDENCY SURVEY DATA*

1. INTRODUCTION

Introduction of The Second Law of thermodynamics (the Law of Entropy) to mainstream economics is often attributed to H. Theil and N. Georgescu-Roegen (see [3], [10]). Since then, entropy and other measures of information content have been interpreted in a variety of applications. Recently, linked with the notion of sustainability, the concept of entropy enjoys a renaissance in economic literature. Sustainability, defined by the World Commission on Environment and Development [12] as a call for continued economic expansion without environmental degradation, focuses on issues of how large the economy should be relative to the environment and how to achieve an optimal inter-temporal allocation of resources. Current economic entropy literature includes elements of information theory, complex systems analysis, and environmental economics. However, few economic applications of entropy measures have been published in Polish literature. They include a study of propensity to smoke [2], correlation analysis [4], and application of relative entropy to evaluate expert forecasts [5].

Entropy of a probability distribution can be interpreted as a measure of uncertainty or, alternatively, as a measure of information content In this paper, I return to the original, information theory definition of entropy to evaluate similarities between *a priori* information supplied by the business tendency surveys (that is, expectations), and *a posteriori* information (that is, realizations). The idea is motivated by a paper by Wędrowska [13] who proposes to interpret the information content of a change of structure from its *a priori* to *a posteriori* form as a measure of degree of similarity (or dissimilarity) of structures. In this paper, *a priori* structure is defined by fractions of respondents expressing expectations, and *a posteriori* structure – by fractions of respondents declaring observed changes in economic variables (realizations).

There are two reasons for undertaking such analysis. First, measures of similarity may provide the means to evaluate accuracy of expectations (predictions) in the business tendency survey. The larger the similarity between expectations and realizations, the better predictive power of expectations. Second, the survey on which empirical

* Research underlying this publication was supported by the Warsaw School of Economics Research Grant No 03/E/0023/100. I would like to thank Dr Ewa Wędrowska for her valuable comments on the preliminary draft of this paper, and an anonymous Referee for helpful suggestions.

part of the paper is based defines expectations as changes "expected in the next 3-4 months". Information content may help to identify the actual forecast horizon used by respondents. This result may, in turn, be useful in other formal analyses of expectations, for example establishing appropriate number of lags in econometric models with expectations or defining dependent variables in quantification models.

In section 2, measures of entropy, information content and dissimilarity of structures are introduced. Business tendency survey data are described in section 3. Empirical results are reported in sections 4 (measures of entropy) and 5 (measures of dissimilarity of structures). Section 6 concludes.

2. MEASURES OF ENTROPY, INFORMATION CONTENT AND DISSIMILARITY OF STRUCTURES

Following Wędrowska [13], let us define structure S^n as a vector $S^n = [s_1, s_2, \ldots, s_n]^T \in \mathrm{R}^n$ which elements s_i ($i = 1, 2, \ldots, n$) fulfill two conditions:

$$0 \leqslant \quad s_i \quad \leqslant 1, \tag{1}$$

$$\sum_{i=1}^{n} s_i = 1. \tag{2}$$

Structure S^n is therefore fully described by a vector of fractions (structure elements) summing to a total of 1.

Amount of information provided by a message (that is, its information content) is defined in information theory in relation to the probability that a given message is received from the set of all possible messages: the less probable the message, the more information it carries. On the basis of the elements of S^n it is now possible to define the empirical measure of entropy introduced by C. E. Shannon in his classic paper *A mathematical theory of communication* as

$$H(S^n) = \sum_{i=1}^{n} s_i \log_2 \frac{1}{s_i}. \tag{3}$$

It is worth noting that value of $H(S^n)$ depends only on characteristics of the structure analyzed, that is, its elements s_i.

An important property of $H(S^n)$ as a measure of entropy is that it reaches its maximum value of $H_{max} = \log_2 n$ if all structure elements s_i are equal (that is, $s_1 = s_2 = \ldots = s_n$; see [6], [8]). As $H(S^n)$ approaches its maximum value, differences between structure elements decrease, and for $H(S^n) = H_{max}$, distribution of structure elements becomes uniform. Also, $H(S^n) = H_{min} = 0$ if one of the elements s_i ($i = 1, 2, \ldots, n$) is equal to 1, and all the remaining structure elements are equal to 0 (that is, distribution is concentrated in one element of structure only). Value of $H(S^n)$ can be therefore interpreted as measure of concentration of elements s_i of structure S^n, and

can be used in empirical setting to evaluate information content of a structure. When several structures ordered in time are available, it is also possible to analyze their dynamics. Empirical values and dynamics of entropy measure $H(S^n)$ for expectations and realizations expressed in business tendency surveys are presented in section 4.

In practice, however, not only the degree of uncertainty associated with *a priori* and *a posteriori* structures may be economically interesting but also extent of changes detected between assumed (*a priori*) and observed (*a posteriori*) structures. In order to analyze the size of change between *a priori* structure S_p^n and *a posteriori* structure S_q^n, relative entropy (or Kullback-Leibler divergence; see [8]) is calculated:

$$I\left(S_q^n : S_p^n\right) = \sum_{i=1}^{n} q_i \log \frac{q_i}{p_i}. \tag{4}$$

Relative entropy is also known as information gain; it measures expected amount of "new" information provided by *a posteriori* structure. One of the properties of $I(S_q^n : S_p^n)$ states that it takes its minimum value of zero if both structures are identical (that is, $S_p^n = S_q^n$), and increases with the size of differences between the structures to infinity (see [8], [13]). $I(S_q^n : S_p^n)$ can be interpreted as degree of change between assumed (*a priori*) and observed (*a posteriori*) structures, and therefore serve as measure of dissimilarity of structures: the larger it is, the less similar the structures are.

In empirical setting, it is more convenient to apply a standardized coefficient defined on interval [0, 1] to facilitate interpretations and comparisons. Chomątowski and Sokołowski in [1] introduce a similarity measure to classify data into comparable phases, and employ it to define clusters of industrial production in Poland. They also provide a related dissimilarity measure that can be used to evaluate extent of change from *a priori* to *a posteriori* structure:

$$P\left(S_q^n : S_p^n\right) = 1 - \sum_{i=1}^{n} \min(q_i, p_i). \tag{5}$$

From the properties of structure defined by (1) and (2) it follows that $P(S_q^n : S_p^n) \in [0, 1]$. The lower limit is attained when analyzed structures are identical, that is, $S_p^n = S_q^n$. As dissimilarities between structures increase, value of $P(S_q^n : S_p^n)$ increases towards the upper limit of 1. Empirical values and dynamics of dissimilarity measure $P(S_q^n : S_p^n)$ employed to evaluate similarities between expectations and realizations expressed in business tendency surveys are described in section 5. They are introduced to supplement results obtained on the basis of entropy measure as both methods reflect structure change from its *a priori* to *a posteriori* state.

3. DATA

Empirical part of this paper focuses on evaluation of information content and dissimilarity between expectations and observed realizations, declared by Polish industrial enterprises in business tendency surveys. Since 1986, qualitative business tendency

surveys are conducted by the Research Institute for Economic Development (RIED) at the Warsaw School of Economics. Originally launched for manufacturing industry, currently they cover households, farming sector, exporters, construction industry, and banking sector as well. In the monthly survey addressed to industrial enterprises (see Table 1), respondents are asked to evaluate both current situation (as compared to last month) and expectations for the next 3 – 4 months by assigning them to one of three categories: increase / improvement, no change, or decrease / decline. Aggregated survey results are regularly published in RIED bulletins (see [7]).

Table 1

Monthly RIED questionnaire in industry

		Observed within last month	Expected for next 3 – 4 months
01	Level of production (value or physical units)	up unchanged down	will increase will remain unchanged will decrease
02	Level of orders	up normal down	will increase will remain normal will decrease
03	Level of export orders	up normal down not applicable	will increase will remain normal will decrease not applicable
04	Stocks of finished goods	up unchanged down	will increase will remain unchanged will decrease
05	Prices of goods produced	up unchanged down	will increase will remain unchanged will decrease
06	Level of employment	up unchanged down	will increase will remain unchanged will decrease
07	Financial standing	improved unchanged deteriorated	will improve will remain unchanged will deteriorate
08	General situation of the economy regardless of situation in your sector and enterprise	improved unchanged deteriorated	will improve will remain unchanged will deteriorate

Source: the RIED database

For empirical analysis, four survey questions have been selected, namely, those pertaining to changes in production, prices, employment and general business conditions. These variables have been analyzed previously, and they have precisely defined counterparts in official statistics necessary for purposes of quantitative analysis (see [11]). *A priori* structure is defined as percentages of respondents who expect increase / no change / decline, and *a posteriori* structure as percentages of respondents who observe increase / no change / decline three and four months later. This definition fulfills conditions (1) and (2). All variables are analyzed for three ownership types: public,

private and total. Since forecast horizon is not precisely defined, both alternatives ($k = 3$ and $k = 4$) are analyzed when calculating measure of dissimilarity $P(S_q^n : S_p^n)$.

Available data cover 161 observations from March 1997 to July 2010. However, since expectations have to be matched with observed realizations to calculate the measure of dissimilarity $P(S_q^n : S_p^n)$, length of time series is reduced either by three (for 3-month forecast horizon) or by four observations (for 4-month forecast horizon). For the purposes of clarity of presentation, all results are reported for the core time period of June 1997 to April 2010 (155 observations).

4. EMPIRICAL RESULTS: ENTROPY

Table 2 presents summary statistics for entropy measure $H(S^n)$ given by formula (3), calculated for all four variables, separately for expectations and observed changes, across ownership sectors [1].

To facilitate comparisons, Figures 1-4 present values of entropy measure $H(S^n)$ for expectations and realizations (for all firms only, to make graphs more legible) against the same scale.

Table 2

Summary statistics for entropy measures

	production						
	expectations				realizations		
	all	public	private		all	public	private
min	1.2650	1.1956	1.2851		1.3289	1.2638	1.2790
max	1.5515	1.5773	1.5679		1.5702	1.5803	1.5751
avg	1.4680	1.4455	1.4741		1.5054	1.4913	1.5093
median	1.4736	1.4588	1.4832		1.5106	1.5079	1.5212
std dev	0.0530	0.0738	0.0583		0.0444	0.0604	0.0485
	prices						
	expectations				realizations		
	all	public	private		all	public	private
min	0.7713	0.6295	0.7566		0.7674	0.5938	0.8155
max	1.3140	1.3355	1.2962		1.2962	1.4447	1.2882
avg	0.9916	0.9755	0.9970		1.0301	1.0224	1.0306
median	0.9821	0.9555	0.9933		1.0272	0.9976	1.0316
std dev	0.1065	0.1331	0.1118		0.1042	0.1630	0.1035

[1] Detailed results are available from author upon request.

c.d. Table 2

	employment						
	expectations				realizations		
	all	public	private		all	public	private
min	0.9398	0.8669	0.8676		1.0486	0.8241	1.0404
max	1.3822	1.4246	1.4101		1.3940	1.4194	1.3901
avg	1.1884	1.1678	1.1925		1.2434	1.2065	1.2605
median	1.1939	1.1723	1.1981		1.2485	1.2228	1.2564
std dev	0.0751	0.0934	0.0895		0.0696	0.1063	0.0713
	general business conditions						
	expectations				realizations		
	all	public	private		all	public	private
min	0.8663	0.8703	0.8778		0.6238	0.7749	0.5354
max	1.4407	1.4902	1.4796		1.3857	1.4478	1.4308
avg	1.2921	1.2575	1.3151		1.1741	1.1334	1.2018
median	1.3020	1.2706	1.3303		1.1929	1.1566	1.2319
std dev	0.0827	0.1080	0.0885		0.1116	0.1138	0.1295

Source: own calculations on the basis of RIED data

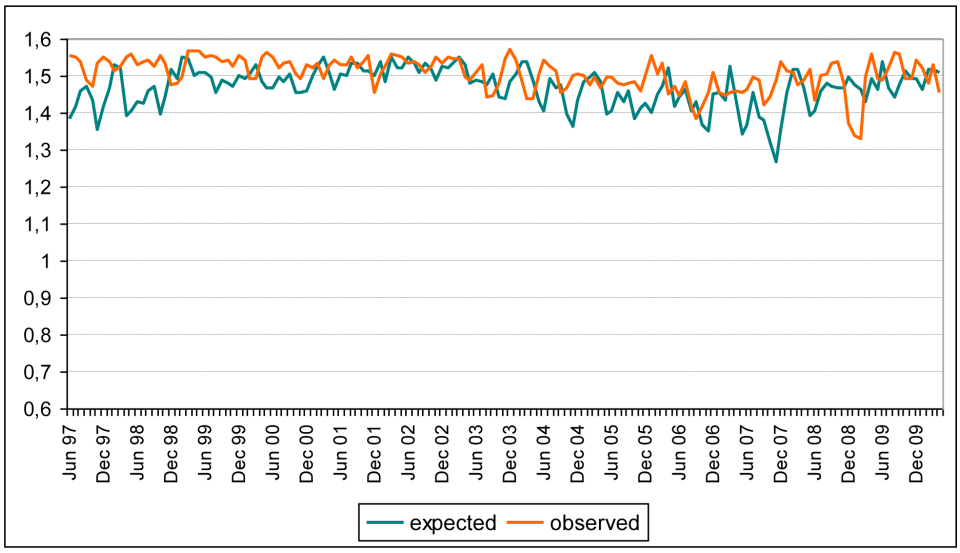


Figure 1. Entropy of production (all firms)

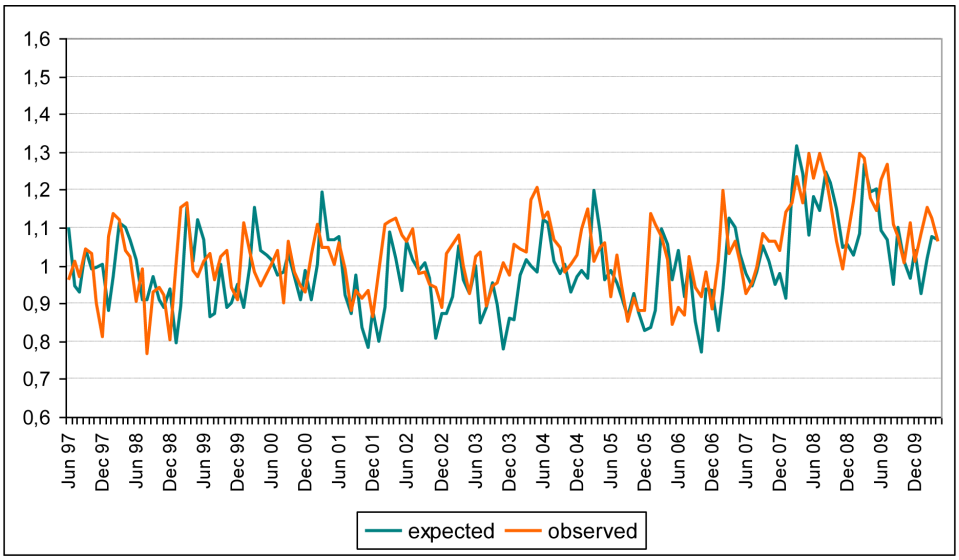


Figure 2. Entropy of prices (all firms)

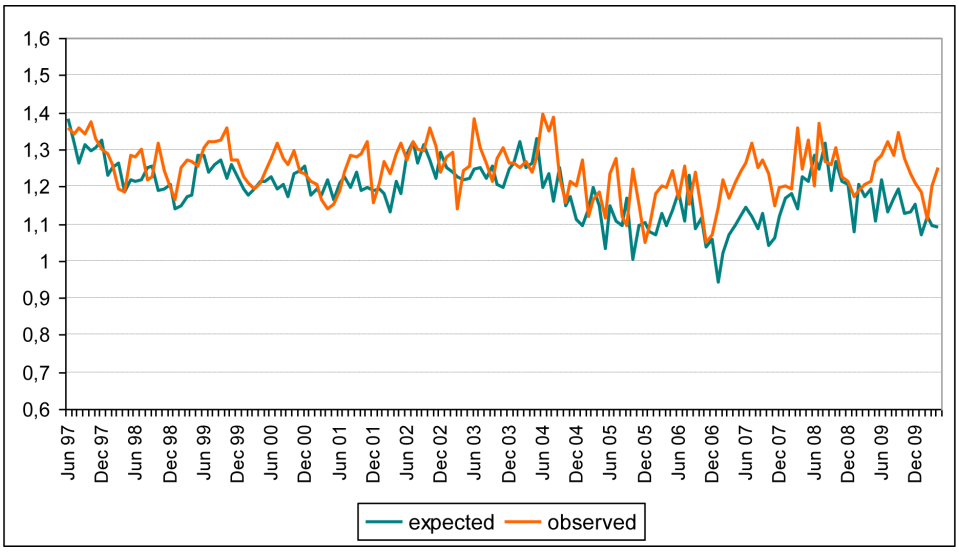


Figure 3. Entropy of employment (all firms)

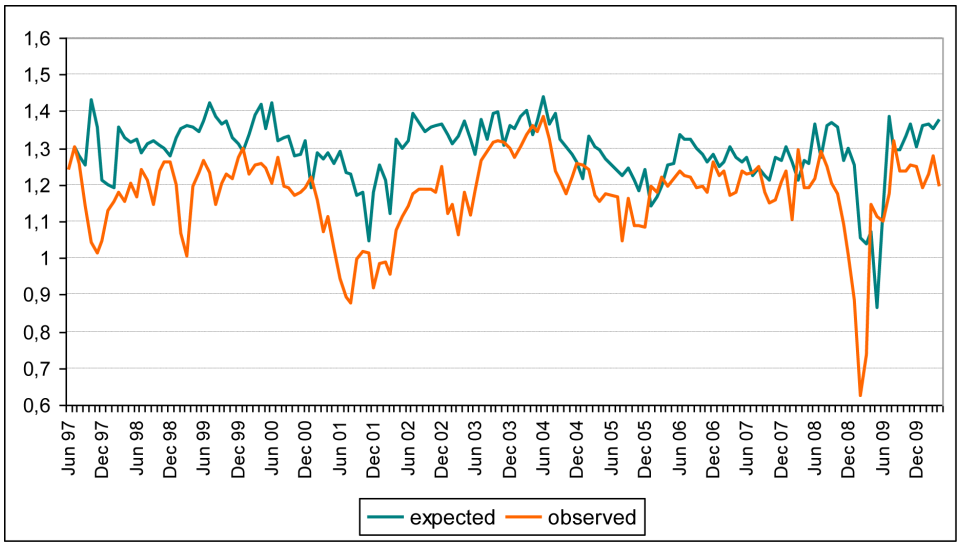


Figure 4. Entropy of general business conditions (all firms)

Source: own calculations on the basis of RIED data

High values obtained in case of production, both in comparison to other variables and in absolute terms, seem to be the most striking result. The maximum value of measure of entropy is $H_{max} = \log_2 n = \log_2 3 = 1.5850$; the closer empirical entropy of a structure to its maximum value, the more uniform the structure is, and therefore the less informative *a priori* structure becomes in relation to *a posteriori* structure. Maximum values obtained for production (1.5515 for expectations and 1.5702 for realizations) are close to the upper limit, and average values (1.4680 for expectations and 1.5054 for realizations) are also very high in comparison to entropy of prices, employment and general business conditions. On the other hand, entropy is equal to zero if one of the elements of a structure is equal to 1, that is, there is no uncertainty associated with distribution of outcomes. Value of zero is not attained for any of the variables analyzed, but the lowest values are observed for prices and realizations of general business conditions.

For production and unemployment, behavior of expected and observed series is similar; in case of employment, expectations exhibit lower averages and medians, across all ownership sectors, than realizations; in case of general business conditions, the opposite is true. Public enterprises exhibit lower average and higher variability of entropy, as measured by standard deviation, than private enterprises, with the sole exception of standard deviation for observed general business conditions.

5. EMPIRICAL RESULTS: DISSIMILARITY OF STRUCTURES

Figures 5-8 provide graphical summary of results obtained for differences between measures of dissimilarity of structures $P(S_q^n : S_p^n)$ calculated for 3-month horizon and 4-month horizon[2].

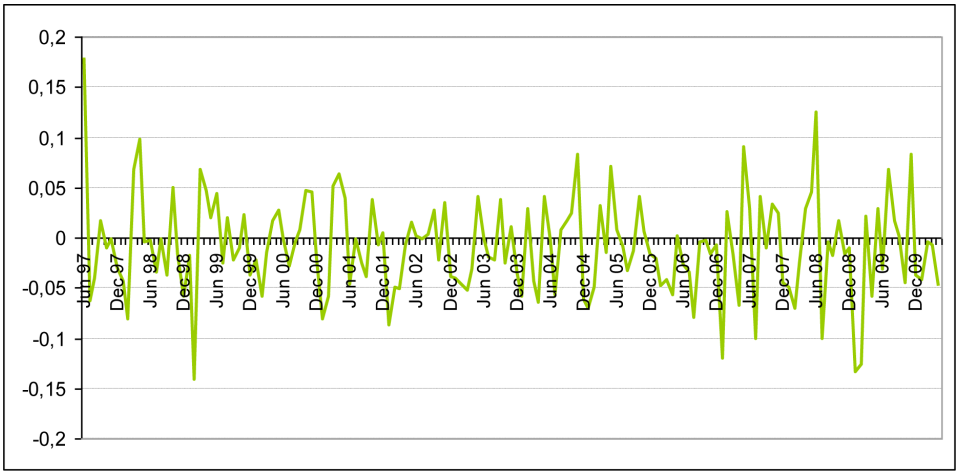


Figure 5. Differences between dissimilarity measures for $k = 3$ and $k = 4$, production (all firms)

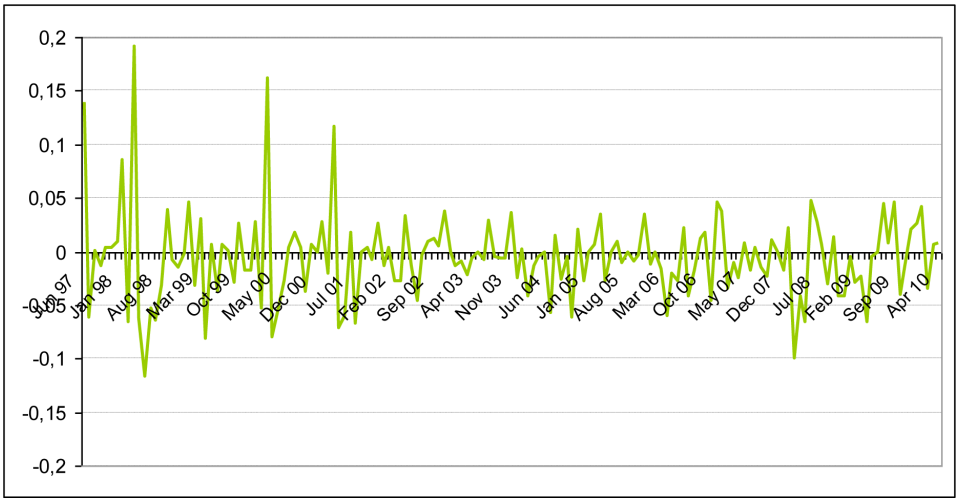


Figure 6. Differences between dissimilarity measures for $k = 3$ and $k = 4$, prices (all firms)

[2] Detailed results (values calculated on the basis of formula (5) for all four variables, three ownership types and separately for 3-month forecast horizon and 4-month forecast horizon) are available from author upon request.

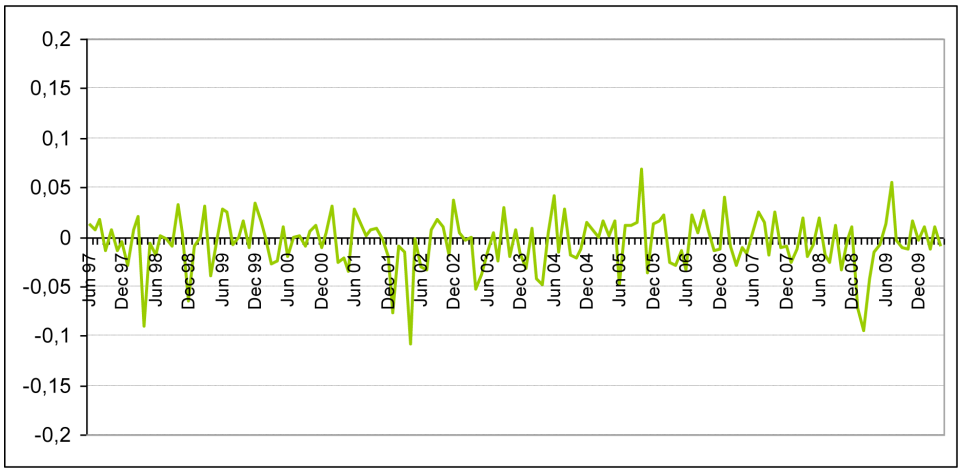


Figure 7. Differences between dissimilarity measures for $k = 3$ and $k = 4$, employment (all firms)

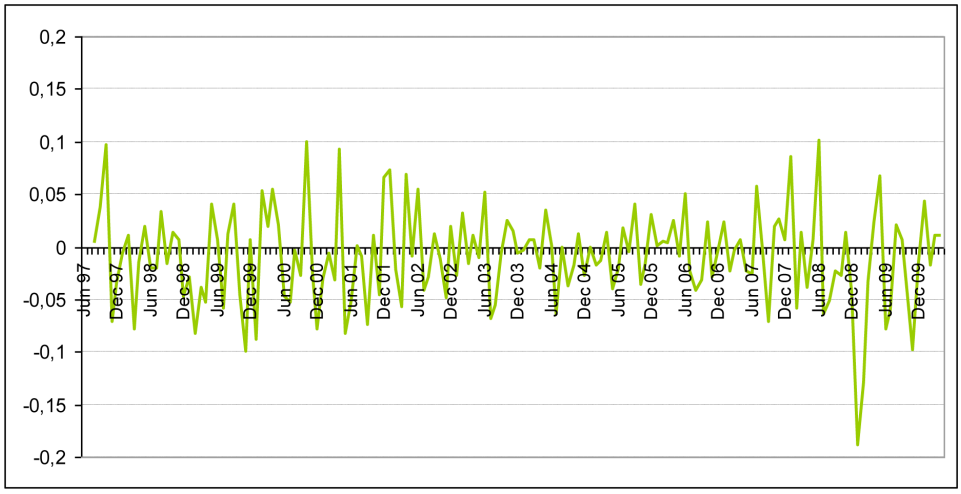


Figure 8. Differences between dissimilarity measures for $k = 3$ and $k = 4$, general business conditions (all firms)

Source: own calculations on the basis of RIED data

In Figures 5-8, negative numbers signify that value of $P(S_q^n : S_p^n)$ for $k = 3$ is smaller than for $k = 4$, that is, structures of realizations and expectations for $k = 3$ are more similar than structures of realizations and expectations for $k = 4$. This can be interpreted as 4-month expectations being less congruent with later observed realization than expectations interpreted in 3-month horizon. This result can be at least partly explained by the well-known observation that uncertainty increases with the forecast horizon, and the tendency of business enterprises to evaluate their performance in quarterly (that is, 3 month) terms.

Table 3 presents summary statistics for differences between dissimilarity measures for $k = 3$ and $k = 4$.

Table 3

Summary statistics for dissimilarity measures: difference between $k = 3$ and $k = 4$

	production	**prices**	**employment**	**business**
min	-0.1420	-0.1160	-0.1080	-0.1880
max	0.1790	0.1920	0.0680	0.1010
avg	-0.0095	-0.0056	-0.0055	-0.0095
median	-0.0100	-0.0050	-0.0030	-0.0060
std dev	0.0480	0.0412	0.0263	0.0446

Source: own calculations on the basis of RIED data

The graphs do not provide basis for determining which forecast horizon is supported by measures of dissimilarity of structure. Average values of dissimilarity statistics (see Table 3) are slightly negative for all variables which suggests that predictions made at the 3-month forecast horizon are more similar to the realizations that those made 4 months in advance. Median values confirm this result. However, the differences between the 3-month and 4-month horizons are very small and should not be considered as a proof of superiority of the shorter forecast horizon, especially in the light of relatively high standard deviations. It is also worth noting that maximum values of dissimilarity measures comparing 3-month and 4-month forecasts for production and prices are twice as high as those obtained for employment and general business conditions. It follows that in production and prices there are (few) periods in which the 4-month forecast horizon is considerably more similar to realizations, although the average remains negative, suggesting that on average the 3-month forecast horizon is closer to the observed changes in analyzed variables.

6. CONCLUDING COMMENTS

On the basis of empirical analysis of the business tendency survey data, the following conclusions have been reached through application of measures of entropy:

1. In case of production, distribution of increase / no change / decrease fractions is relatively uniform, leading to high entropy and providing little information.
2. Entropy of prices is relatively low; since value of entropy allows to evaluate degree of concentration, in case of prices fractions of survey answers seems to be particularly centered on one of the three options provided in the questionnaire. In theory, answers might be centered on either of the three options (increase / no change / decrease) and vary from one questionnaire to another. In practice, however, they are heavily biased towards the "no change" category. In the analyzed period (that is, March 1997 – July 2010), no change in prices is always expected by the majority of

the respondents – that is, "no change" fraction constantly remains the largest among the three options. This result does not hold in case of production, employment, or general business conditions expectations, in confirmation of the results of entropy analysis.

3. Entropy of general business conditions exhibits the highest variability which may be interpreted as volatile changes in information content of surveys from one month to another. In contrast, entropy of production is the least variable.
4. Generally, public enterprises exhibit lower entropy (as measured by average) and higher variability (as measured by standard deviation) than private enterprises; that is, for public enterprises concentration of answers to the survey questions is higher and also more variable.

To evaluate whether structures for 3-month or 4-month forecast horizons are more similar to observed realizations, measures of dissimilarity of structures were calculated. Unfortunately the results do not provide clear answer to this question. A slight tendency towards the 3-month forecast horizon is noted on the basis of negative (but very small in absolute terms) average values of dissimilarity measures across all four variables.

To summarize, results obtained on the basis of entropy and dissimilarity measures provide new insights into behavior of expectations and realizations expressed in business tendency surveys. As this is the first attempt to empirically address the question of information content of the RIED survey data, more work is clearly needed. One of the issues that merit further analysis is whether current situation of an enterprise systematically influences its expectations, and consequently degree of concentration of answers on a particular option. To evaluate usefulness of entropy measures in analyzing questionnaire data, predictive value of a *priori* information should be studied, and predictive properties of various statistic tools compared for different distributions of answers.

Author is employed at the Institute of Econometrics, Warsaw School of Economics.

REFERENCES

[1] Chomątowski S., Sokołowski A. (1978), *Taksonomia struktur*, Przegląd Statystyczny 2:217-226.
[2] Doszyń M. (2002) *Skłonności a entropia*, Przegląd Statystyczny 49:73-78.
[3] Georgescu-Roegen N. (1971) *The Entropy Law and the Economic Process*, Harvard University Press, Cambridge.
[4] Kempa W. (2002), *Zastosowanie entropii empirycznej w badaniu związku korelacyjnego dwóch cech* Przegląd Statystyczny 49:163-173.
[5] Kowalczyk H. (2010), *O eksperckich ocenach niepewności w ankietach makroekonomicznych*, Bank i Kredyt 5:101-122.
[6] Przybyszewski R., Wędrowska E. (2005), *Aksjomatyczna teoria entropii*, Przegląd Statystyczny 52:85-101.
[7] RIED (2010), *Business survey. June 2010*, Warsaw School of Economics, Warsaw.
[8] Rényi A. (1961), *On measures of entropy and information,* Proceedings of the 4^{th} Berkeley Symposium on Mathematics, Statistics and Probability, pp. 547-561.

[9] Shannon C. E. (1948)*A mathematical theory of communication*, The Bell System Technical Journal 27:379-423, 623-656.
[10] Theil H. (1967), *Economics and Information Theory*, North-Holland Publishing Company, Amsterdam.
[11] Tomczyk E. (2005) *Are expectations of Polish industrial enterprises rational? Evidence from business tendency surveys*, in:Adamowicz E., Klimkowska J. (eds.) *Economic Tendency Surveys and Cyclical Indicators. Polish contribution to the 27th CIRET Conference,* Warsaw School of Economics, Warszawa.
[12] WCED (1987), *Report of the World Commission on Environment and Development: Our Common Future*, NGO Committee on Education (http://www.un-documents.net/wced-ocf.htm).
[13] Wędrowska E. (2009), *Oczekiwana ilość informacji o zmianie struktur jako miara niepodobieństwa struktur*, paper presented at the XIth Conference "Dynamic Econometric Models", September 2009, Toruń.

APPLICATION OF MEASURES OF ENTROPY, INFORMATION CONTENT AND DISSIMILARITY OF STRUCTURES TO BUSINESS TENDENCY SURVEY DATA

Summary

This paper evaluates similarities between *a priori* information supplied by business tendency surveys (that is, expectations), and *a posteriori* information (that is, realizations). A *priori* structure is defined by fractions of respondents expressing expectations, and *a posteriori* structure – by fractions of respondents declaring observed changes in economic variables (realizations). On the basis of empirical analysis of the business tendency survey data on production, prices, employment and general business conditions, the following conclusions have been reached. Production time series exhibits the highest entropy, and prices data – the lowest. Since value of entropy allows to evaluate degree of concentration, in case of prices fractions of survey answers seems to be particularly centered on one of the three options provided in the questionnaire (that is, increase – no change – decrease). Entropy of general business conditions exhibits the highest variability which may be interpreted as volatile changes in information content of surveys from one month to another; in contrast, entropy of production is the least variable. It is also found that public enterprises exhibit lower entropy (as measured by average) and higher variability (as measured by standard deviation) than private enterprises.

Key words: tendency surveys, expectations, entropy, dissimilarity of structures

ZASTOSOWANIE MIAR ENTROPII, ZAWARTOŚCI INFORMACYJNEJ I NIEPODOBIEŃSTWA STRUKTUR DO DANYCH TESTU KONIUNKTURY

Streszczenie

Artykuł bada podobieństwo między informacją *a priori* dostarczaną przez respondentów testu koniunktury (oczekiwaniami) a informacją *a posteriori* (zaobserwowanymi realizacjami). Struktura *a priori*

definiowana jest poprzez odsetki respondentów wyrażających swoje oczekiwania, a struktura *a posteriori* – przez odsetki respondentów stwierdzających zaobserwowane zmiany. Na podstawie empirycznej analizy danych testu koniunktury na temat produkcji, cen, zatrudnienia oraz ogólnej sytuacji gospodarczej, sformułowano następujące wnioski. Produkcja cechuje się najwyższą entropią, a ceny – najniższą. Ponieważ poziom entropii może być interpretowany jako stopień koncentracji, w przypadku cen odsetki odpowiedzi na pytania testu koniunktury wydaje się być szczególnie mocno skoncentrowany na jednej z trzech opcji (wzrost – brak zmiany – spadek). Entropia ogólnej sytuacji gospodarczej wykazuje największą zmienność, co można zinterpretować jako przejaw dynamicznych zmian zawartości informacyjnej ankiety w poszczególnych miesiącach; entropia produkcji jest najbardziej stabilna. Co więcej, przedsiębiorstwa sektora publicznego cechuje średnio niższa entropia i wyższa jej zmienność (mierzona odchyleniem standardowym) niż przedsiębiorstwa prywatne.

Słowa kluczowe: badania ankietowe, oczekiwania, entropia, niepodobieństwo struktur

PRZEGLĄD STATYSTYCZNY
R. LVIII – ZESZYT 1-2 – 2011

MICHAŁ BRZEZIŃSKI

STATISTICAL INFERENCE ON CHANGES IN INCOME POLARIZATION IN POLAND

1. Introduction

Over the last years, measurement and estimation of economic polarization has been getting an increased attention among economists and other social scientists. Starting with contributions of Esteban and Ray [14] and Wolfson [35], several researchers proposed different measures of polarization, often developed formally in an axiomatic framework[1]. Major applications of these and other approaches to polarization measurement include analysis for the UK [25], Spain [21, 22], China [36], Uruguay [23], Poland [27], and Denmark [24]. There are also studies devoted to the regional polarization in Russia [20], the European Union [18] and Central and Eastern European countries [19], as well as many papers conducting cross-country analysis [8, 10, 11, 13, 28, 32].

Economic polarization as conceptualized in these studies concerns the distribution of socio-economic characteristics (for example income, consumption, race, education) and refers to the extent to which the society is grouped in a number of clusters. In other words, in a polarized society incomes or other relevant attributes become concentrated around two or more diverging poles. Such a phenomenon is related, but certainly conceptually different from the more familiar concept of inequality. For example, it is intuitively easy to imagine a society converging in incomes to two local poles (for example grouping 'the rich' and 'the poor'). Such a society would display lowering level of inequality as measured by standard inequality indices, since the overall income dispersion would decrease, but it would also become more clustered around the two poles and therefore more bi-polarized[2].

What gives polarization a special appeal, both from the abstract and policy-oriented perspective, is that it has been theoretically linked to the phenomenon of social conflicts. Esteban and Ray [15] argued that it is polarization, and not inequality as commonly perceived, which is correlated with social conflicts like organised, large-scale

[1] The major contributions include Wang and Tsui [34], Chakravarty and Majumder [5], Zhang and Kanbur [36], Duclos et al. [8], and Esteban et al. [13].

[2] The issue of whether polarization and inequality are distinguishable empirically has been a matter of some debate. Ravallion and Chen [31] and Zhang and Kanbur [36] argued that measures of polarization generally do not generate very different results from those of standard measures of inequality. Esteban [12], Duclos et al. [8], and Lasso de la Vega and Urrutia [28] provide evidence that the two sets of indices significantly differ empirically.

protests, strikes, demonstrations, or even revolts and armed unrests. They showed that in the case of bipolarized society the level (or intensity) of social conflicts increases with the magnitude of polarization[3].

From another point of view, polarization has often been associated with the 'disappearing of the middle class' – a phenomenon characteristic for the US and the UK in the 1980s [25, 35]. Indeed, if incomes concentrate around two opposite distributive poles, then the size of the middle class has to decrease. The common opinion and a number of economic arguments suggest that a stable and sizable middle class is a necessary condition for successful economic and political transition and development (see, e.g., [2], pp. 3–4). Therefore, high or increasing level of bi-polarization can pose a serious problem for policymakers.

Recently, Kot [27] has offered an important contribution to both theoretical and empirical literature on economic polarization. He introduced a new class of polarization indices based on the concept of the Lorenz curve. Using Polish Household Budget Survey (HBS) data, he estimated these new indices as well as several existing polarization measures, including those introduced in Wolfson [35] and Esteban et al. [13], for the distribution of household expenditures in Poland during 1993-1997 and 1998-2005[4]. He found that during the first period there was no clear trend in the economic polarization, while in the second period polarization has increased by 0.85 to 8.1%, depending on the polarization index used.

This paper attempts to make two contributions. First, we complement Kot's [27] empirical analysis by estimating another important polarization index proposed by Duclos et al. [8] for the income distribution in Poland during 1998-2007. Beside developing the axiomatic foundation for the index, Duclos et al. [8] have provided also a formula for an asymptotic variance for their polarization index estimated from household survey samples. It is therefore possible to conduct formal statistical inference on this index by calculating confidence intervals and testing hypotheses about the changes in the values of the index over time. Hence, the second contribution of the paper is to test whether estimated changes in income polarization are statistically significant. It allows us to make polarization comparisons for Poland robust to sampling variability and to formulate reliable conclusions on whether income polarization has increased or not.

Section 2 of the paper introduces polarization measure as well as the methods of its sample estimation and tools of statistical inference as offered by Duclos et al. [8]. Micro data from Polish Household Budget Survey study are discussed in section 3. Section 4 reports and discusses empirical results, while section 5 concludes.

---

[3] See also Esteban and Ray [16, 17]. Esteban and Ray ([16], p. 164) argue that empirical evidence on the connection between *inequality* and social conflicts is ambiguous and inconclusive.

[4] The period between 1993 and 2005 is divided into two sub-periods because HBS data for these sub-periods are not directly comparable (see also section 3).

2. MEASURES OF POLARIZATION

The approach to measuring economic polarization introduced in Duclos et al. [8] has two attractive features – one theoretical, the other practical. The theoretical advantage of this formulation is that it is formulated in the so-called identification-alienation framework, introduced in Esteban and Ray [14]. This framework allows for formal analysis of the abovementioned link between polarization and the phenomenon of social conflict [15, 16, 17]. The practical benefit of the approach is that it is carefully operationalized in a statistical framework. Duclos et al.'s [8] methodology offers distribution-free statistical inference procedures, which allows to check whether the resulting estimates are not the effect of sampling noise and measurement errors[5].

Identification-alienation framework, introduced in Esteban and Ray [14], suggests that polarization can be measured as the effect of two interrelated mechanisms: (1) alienation, which is felt by individuals from a given group (defined by income class, religion, race, education etc.) toward individuals belonging to other groups, and (2) identification, which unites members of any given group. This approach assumes that polarization requires that individuals identify with other members of their socio-economic group and feel alienation to members of other groups. The approach of Duclos et al. [8] is proposed for the space of distributions of incomes (or consumption expenditures, wealth, etc.) described by income density functions (f), with integrals of these functions corresponding to various population sizes. It is assumed that identification at income y depends on the density at x, $f(x)$, while alienation between two individuals with incomes y and x is given by $|y-x|$. By imposing a set of axioms on the identification-alienation structure, Duclos et al. [8] derive the following polarization measure (DER index):

$$P_a(f) = \iint^{1+\alpha} f(y)\,|y-x|\,dydx, \tag{1}$$

where α is an ethical parameter expressing the weight given to the identification part of the framework. The axioms introduced by Duclos et al. [8] require that α must be bounded in the following way: $0.25 \leqslant \alpha \leqslant 1$. When $\alpha = 0$, the DER index is equal to the twice the popular Gini coefficient of inequality, which for density f and all incomes normalized by their mean takes the form

$$G(f) = \frac{1}{2}\iint f(y)\,|y-x|\,dydx. \tag{2}$$

Therefore, the lowest admissible value for DER index of $\alpha = 0.25$ should produce polarization results close to those of inequality measured by Gini index, while $\alpha = 1$ leads potentially to the highest disparity between the Gini coefficient and the DER index[6].

[5] See also Duclos et al. [9] for the more complete presentation of their statistical framework.

[6] In order to make comparisons between the Gini index and the DER index easier, in our empirical application in section IV we compute all measures for incomes normalized by the mean income and divide DER indices by 2.

Duclos et al. [8] analyze also the problem of the statistical estimation of their index. It can be shown that for every income distribution function F with associated density function f and mean μ, Eq. 1 can be restated in the form

$$P_\alpha(F) = \int_y f(y)^\alpha \alpha(y)\, dF(y), \tag{3}$$

with $a(y) \equiv \mu + y(2F(y) - 1) - 2\int_{-\infty}^{y} xdF(x)$.

Suppose next that the index is to be estimated using a random sample of n i.i.d (independent and identically distributed) sample observations of income yi, i =1,..., n, drawn from the distribution $F(y)$ and ordered such that $y_1 \leqslant y_2 \leqslant ... \leqslant y_n$. It is also assumed that a set of sampling weights is provided with w_i being a sampling weight on observation i and $\bar{w} = \sum_{j=1}^{n} w_j$ being the sum of weights. All incomes are normalized by the sample weighted mean income $\left(\mu = \sum_{i=1}^{n} w_i y_i / \sum_{i=1}^{n} w_i\right)$. Using Eq. 3, the DER index can be then estimated by

$$P_\alpha\left(\hat{F}\right) = \bar{w}^{-1} \sum_{i=1}^{n} w_i \hat{f}(y_i)^\alpha \hat{a}(y_i), \tag{4}$$

where $\hat{a}(y_i)$ is given as

$$\hat{a}(y_i) = \mu + y_i\left(\bar{w}^{-1}\left(2\sum_{j=1}^{i} w_j - w_i\right) - 1\right) - \bar{w}^{-1}\left(2\sum_{j=1}^{i-1} w_j y_j + w_i y_i\right), \tag{5}$$

and where $\hat{f}(y_i)^\alpha$ is estimated non-parametrically using kernel density estimation methods[7]. Duclos et al. [8] use a symmetric weighted kernel function $K(u)$, defined such that $\int_{-\infty}^{+\infty} K(u)du = 1$ and K(u) $\geqslant$ 0. The estimator $\hat{f}(y)$ is then defined as

$$\hat{f}(y) \equiv (\bar{w}h)^{-1} \sum_{i=1}^{n} w_i K\left(\frac{y - y_i}{h}\right), \tag{6}$$

where the smoothing parameter h, which is usually called the bandwidth, describes the width of the density window around each point, and kernel function $K(u)$ is specified as the Gaussian kernel, defined by

$$K(u) = (2\pi)^{-0.5} \exp^{-0.5u^2}. \tag{7}$$

[7] Kernel density estimation methods are presented in detail in, for example, Pagan and Ullah ([29], cha. 2) and Silverman ([33], cha. 3). See also Jenkins [25], and Duclos and Araar ([7], cha. 15) for analysis of these methods in the context of income distribution.

The optimal value of the bandwidth ($h*$) is chosen as to minimize the mean square error (MSE) of the estimator (1) for a given sample size (n). Duclos et al. [8, 9] devise computational formulas, which approximate h^*. For a normal distribution with variance and a Gaussian kernel function used in (3), the formula is

$$h^* \cong 4.7 n^{-0.5} \sigma \alpha^{0.1}, \tag{8}$$

where α is again polarization sensitivity parameter. For an income distribution characterized by skewness greater than about 6, a better approximation is given by a more complex formula of the form

$$h^* \cong n^{-0.5} IQ \frac{(3.76 + 14.7\sigma_{\ln})}{(1 + 1.09 \cdot 10^{-4}\sigma_{\ln})^{7268+15323\alpha)}}, \tag{9}$$

where IQ is the interquartile range (the difference between the 75^{th} and the 25th percentile) and is the variance of the logarithms of income[8].

Duclos et al. [8, 9] further show that under standard regularity conditions concerning the finiteness of moments of certain variables, if h in Eq. 6 vanishes when n tends to infinity, then $n^{0.5}\left(P_\alpha\left(\widehat{F}\right) - P_\alpha(F)\right)$ has an asymptotic limiting normal distribution $N(0, V_\alpha)$, with

$$V_\alpha = var_{f(y)}\left((1+\alpha) f(y)^\alpha a(y) + y \int f(x)^\alpha dF(x) + 2 \int_y^\infty (x - y) f(x)^\alpha dF(x)\right). \tag{10}$$

Variance given by Eq. 10 is distribution-free in the sense that it can be computed without knowledge of the form or the parameters of the distribution from which the sample is drawn. The expression (11) is operationalized by the use of the delta method (see, e.g., Rao [30], pp. 385-391) in the DASP software package written for STATA [1, 7]. In this paper, asymptotic variance and standard errors for the DER index are computed with the help of the DASP package.

Using estimates of the DER index and its asymptotic variance, it is possible to construct confidence intervals for the index and to test hypotheses about the changes in its value. The statistical inference in this paper is based on the asymptotic t-type statistic computed using Eq. 3 and the variance estimate of the DER index computed on the basis of Eq. 10. As we are mainly interested in changes in income polarization in time, the relevant hypothesis states that two different income distributions have the same value of the DER index. Assuming that we have independent samples drawn from two distributions (e.g. income distributions in periods t_0 and t_1), we can compute estimates $P_\alpha\left(\hat{F}_{t_0}\right)$ and $P_\alpha\left(\hat{F}_{t_1}\right)$ from the two samples using formula (3), and variance estimates $\hat{V}_\alpha(t_0)$ and $\hat{V}_\alpha(t_1)$ using Eq. 10[9]. For a null hypothesis that $P_\alpha(F_{t_0}) = P_\alpha(= F_{t_1})$, a

[8] The skewness of income distribution estimated from household survey data may often be greater than 6 (cf. Kot [27], p. 78), especially if extreme incomes are not eliminated, recoded or dealt with in other way.

[9] Cf. Davidson and Flachaire [6] for asymptotic inference for differences in inequality indices.

t-type statistic is

$$W = \frac{P_\alpha\left(\hat{F}_{t_0}\right) - P_\alpha\left(\hat{F}_{t_1}\right)}{\left(\hat{V}_\alpha\left(t_0\right) + \hat{V}_\alpha\left(t_1\right)\right)^{0.5}}. \tag{11}$$

Finally, using expression (11) we compute asymptotic P values based on the standard normal distribution or on the Student distribution with n degrees of freedom.

3. DATA

This paper uses data from Household Budget Survey (HBS) study conducted yearly by the Polish Central Statistical Office (CSO). We use yearly HBS micro-data for the period 1998-2007. An important modification of the HBS occurred experimentally in 1997, and definitively in 1998, when in order to fulfil Eurostat recommendations, new definitions of some core concepts (i.e. disposable income) were implemented. Due to this change it is rather difficult to construct fully comparable series of household disposable incomes for the period before 1998 and after this year. Therefore, in this paper we use HBS data from 1998 to 2007 (the last available year). The HBS sample design has changed several time during the period under study[10]. From the point of view of the present paper, it is important to notice that during 1996-2000 the design assumed sampling of primary sampling units (PSU's) for a period of four years, while since 2001 every year a new subsample is drawn for use during two years. Therefore, year-to-year samples are correlated and the statistic W given by Eq. 11 should take into account also the covariance of the two estimated polarization indices. To avoid this complication we study statistical inference on the DER index using independent samples for the years 1998, 2001, 2004 and 2007.

Household net disposable income (i.e. post-tax-and-transfer income) is the main income concept used. It includes cash wages and salaries, self-employment income (including farm income), cash property income, social transfers (including social insurance, social assistance) and other income. Income taxes, mandatory payroll taxes and gifts donated to other households are not included. As it is standard in income distribution literature, we consider the individual as the main unit of analysis. In order to obtain personal income distributions, all household observations are weighted by the product of household weights provided in the HBS and household size[11]. All incomes are divided by equivalence scales defined as $h^{0.5}$, where h is household size, to adjust for the size and composition of households. Finally, to obtain real equivalent disposable income we use CPI deflator to express all incomes in December 2007 price levels.

[10] The detailed description of the HBS design and its other features can be found in Kordos et al. [26] and Central Statistical Office [4].

[11] HBS weights are non-response weights adjusting sample data for the differential non-response rates of different types of households. The method of estimating these weights has changed several times between 1998 and 2004. See Kordos et al. [26] and Central Statistical Office [4] for details.

4. RESULTS

Figure 1 shows trends in the values of the DER index for $\alpha \in \{0.25, 0.5, 0.75, 1\}$ during the period under study. To verify whether income polarization as measured by the DER indices is indeed behaving differently from income inequality as measured by the Gini index, which is equal to the DER index with $\alpha = 0$, the changes in the latter are presented as well. As expected, changes in the DER index with $\alpha = 0.25$ are very similar in quantity and direction to the changes in the Gini index. The same can be said for other values of the parameter α except $\alpha = 1$. It can be observed that in this case income polarization as measured by the DER index is changing in a markedly different ways from inequality. For example, while the Gini index is clearly decreasing over 2004-2006, DER (α =1) is slightly increasing. Such opposite tendencies are displayed by the indices for 2002-2003 and 1998-1999 periods as well. Our analysis confirms, therefore, the conclusion of Duclos et al. [8] that the DER index, at least for the highest admissible value of α, is empirically distinct from the Gini index.

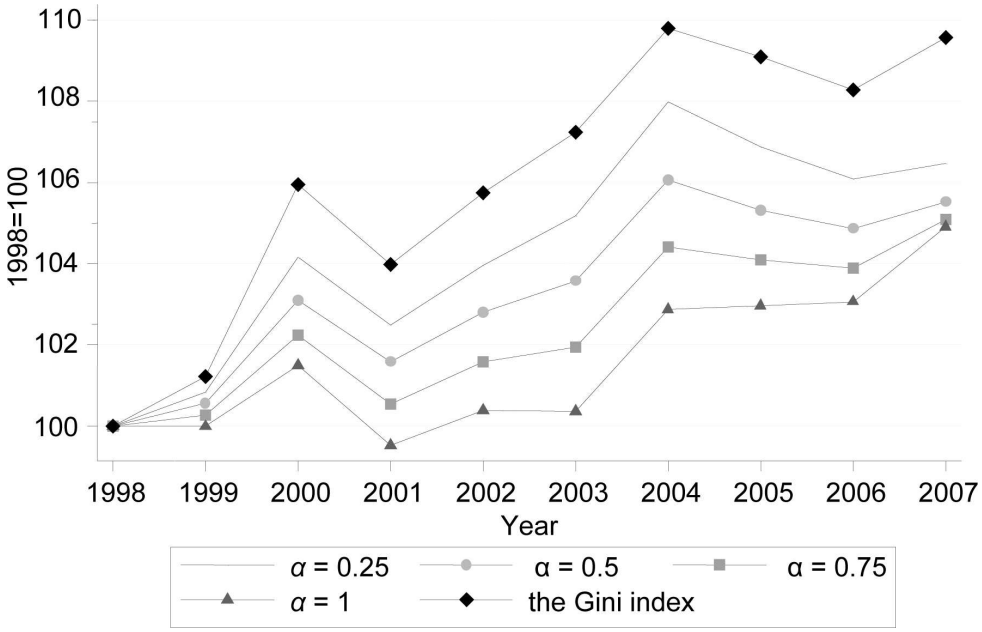


Figure 1. Changes in income polarization (DER index) in Poland, 1998-2007

Throughout the period under study, there was an increase in income polarization ranging from 4.9 to 6.5%, depending on the value of α parameter. This finding is consistent with the slight increase in the polarization of consumption expenditures over the 1998-2005 period found by Kot [27]. Table 1 reports values of the DER index for $\alpha = 1$, which gives the highest weight to the identification part of the identification-alienation framework in the Duclos et al.'s [8] account. Moreover, for our datasets, this value of the α is the one for which polarization is empirically clearly different from

inequality. Table 1 shows also the values of the standard error for the estimate of the index along with 95% and 99% confidence intervals. For the comparison between 1998 and 2007 both types of confidence intervals are non-overlapping and therefore we may conclude that the difference in income polarization as measured by the DER index with α =1 is statistically significant. Similarly, there is a statistically significant increase in income polarization between 2001 and 2007. However, for 4 out of 6 paired comparisons under study (i.e. 1998 and 2001, 1998 and 2004, 2001 and 2004, 2004 and 2007) both types of confidence intervals are overlapping and therefore they cannot be used to state unambiguously whether income polarization has changed. Hence, to provide statistical inference on these changes we turn to hypotheses testing.

Table 1

Income polarization according to the DER index (α = 1), Poland, 1998-2007

	DER index	SE	95% CI		99% CI	
			LB	UB	LB	UB
1998	0.1614	0.0010	0.1594	0.1635	0.1587	0.1641
1999	0.1614	0.0018	0.1578	0.1650	0.1566	0.1661
2000	0.1638	0.0015	0.1610	0.1667	0.1600	0.1675
2001	0.1606	0.0009	0.1589	0.1623	0.1584	0.1629
2002	0.1620	0.0013	0.1594	0.1646	0.1586	0.1654
2003	0.1620	0.0009	0.1603	0.1637	0.1597	0.1642
2004	0.1660	0.0013	0.1634	0.1687	0.1626	0.1695
2005	0.1662	0.0011	0.1641	0.1683	0.1635	0.1689
2006	0.1664	0.0010	0.1645	0.1682	0.1639	0.1688
2007	0.1693	0.0012	0.1669	0.1717	0.1662	0.1725

Notes: Own calculations using HBS data. Column 1 reports point estimates of the DER index calculated on the basis of the household data with weights equal to the product of the household weight and the size of the household. Column 2 reports standard errors of the DER index. Columns 3 and 4 report lower and upper bounds (LB and UB) for, respectively, 95% and 99% confidence intervals.

Table 2 shows, first, the change (D) in income polarization as measured by the DER index (α = 1) for every pair among the analyzed years (1998, 2001, 2004, 2007) with index for the later year always being the minuend. This is followed by standard error (SE) for the estimate of the measured change in the DER index, and the probability (P value) that such a change is greater than zero (P value = Pr($D > 0$). If D is less than 0, then $1-P$ value = Pr($D < 0$). Among the paired comparisons, the change from 1998 to 2001 is not statistically significant at the conventional 5% level with P value equal to 0.28. The change from 2004 to 2007 is significant only at the 5% level, but not at the 1% level. All other changes are statistically significant at standard levels.

The main conclusion from these results is that short-period observed changes in income polarization in Poland, which are usually quantitatively rather small, can often reflect rather sampling variability than the real changes in the underlying income distribution. Therefore, such short-period changes, including year-to-year changes, should be treated cautiously especially is they are to be used as a basis for policy recommendations. On the other hand, for the HBS data even moderate changes in the DER index, as observed during the 6- and 9-year time spans among our paired comparisons, appear to be statistically significant.

Table 2

Hypotheses tests for changes in the DER index (α = 1)

		2001	2004	2007
1998	*D*	–0.000792	0.004614	0.007921
	SE	0.001362	0.001697	0.001611
	P value	0.2804	0.0033	0.0000
2001	*D*		0.005406	0.008713
	SE		0.001602	0.001511
	*P*value		0.0004	0.0000
2004	*D*			0.003307
	SE			0.001819
	P value			0.0345

Notes: *D* denotes differences in measured values of DER indices over time, SE is the standard error of *D*, and *P* value is the probability that *D* is greater than 0.

5. CONCLUSIONS

This paper has provided estimates of the DER index of economic polarization, calculated on the basis of income distribution HBS data for Poland between 1998 and 2007. It has also offered statistical inference on the changes in income polarization by estimating confidence intervals and testing statistical hypotheses. The main conclusion is that for the entire period under study there was a moderate rise in income polarization in the range of 4.9 to 6.5%, depending on the polarization-sensitivity parameter α. For the period 1998-2007, the changes in the DER index with α = 1, for which polarization is empirically distinguishable from inequality, are statistically significant at conventional levels. The paper has found, however, that for smaller increases, which usually occur over shorter periods, it may be difficult to conclude whether observed changes are due to the changes in the income distribution in the population or due to the sampling variability of the estimates.

Finally, a caveat is in order concerning the methods used to estimate the variance of the DER index from the HBS data. These methods, introduced in Duclos et al. [8], assume that income observations are i.i.d., whereas HBS is a complexly designed survey with stratification, clustering and weighting of sample observations. While weighting can be rather easily incorporated in the statistical inference procedures, the effects

of stratification and clustering are harder to dealt with. However, accounting for the complex sample design would almost certainly increase variance estimates for the DER index[12]. This would lead to wider confidence intervals, larger P values in our empirical application, and could force us to reverse the direction of inference in at least some cases. We leave the issue of accounting for complexity of survey design in inference for polarization measures for future research.

REFERENCES

[1] Araar A. and Duclos J.-Y., *User Manual for Stata Package DASP: Version 2.0*, PEP, CIRPEE, World Bank and WIDER, 2009.

[2] Banerjee A and Duflo E., What is middle class about the middle classes around the world? *Journal of Economic Perspectives*, 22(2) (2008), pp. 3-28.

[3] Bhattacharya D., Inference on Inequality from Household Survey Data, *Journal of Econometrics*, 137(2) (2007), pp. 674-707.

[4] Central Statistical Office, *Household Budget Surveys in 2007*, CSO, Warszawa, 2008.

[5] Chakravarty S.R. and Majumder A., Inequality, polarization and welfare: theory and applications. *Australian Economic Papers*, 40(1) (2001), pp. 1-13.

[6] Davidson R., Flachaire E., Asymptotic and bootstrap inference for inequality and poverty measures, *Journal of Econometrics*, 141(1) (2007), pp. 141-166.

[7] Duclos J.-Y. and Araar A., *Poverty and Equity: Measurement, Policy, and Estimation with DAD*, Springer Publishing Company, New York, 2006.

[8] Duclos J.-Y., Esteban J. and Ray D., Polarization: concepts, measurement, estimation. *Econometrica*, 72(6) (2004), pp. 1737-1772.

[9] Duclos J.-Y., Esteban J. and Ray D., Polarization: concepts, measurement, estimation. in C. Barrett (ed.), *The Social Economics of Poverty: Identities, Groups, Communities and Networks*, Routledge, London, 2005, pp. 56-105.

[10] Duro J.A., Another look to income polarization across countries. *Journal of Policy Modeling*, 27(9) (2005), pp. 1001-1007.

[11] Duro J.A., International income polarization: a note. *Applied Economics Letters*, 12(12) (2005), pp. 759-762.

[12] Esteban J., *Economic polarization in the Mediterranean Basin*, Els Opuscles del CREI #10, Barcelona, 2002.

[13] Esteban J., Gradín C. and Ray D., An extension of a measure of polarization, with an application to the income distribution of five OECD countries. *Journal of Economic Inequality*, 5(1) (2007), pp. 1-19.

[14] Esteban J. and Ray D., On the measurement of polarization. *Econometrica*, 62(4) (1994), pp. 819-851.

[15] Esteban J. and Ray D., Conflict and distribution. *Journal of Economic Theory*, 87(2) (1999), pp. 379-415.

[16] Esteban J. and Ray D., Polarization, fractionalization and conflict. *Journal of Peace Research*, 45(2) (2008), pp. 163-182.

[17] Esteban J., Ray D., Linking conflict to inequality and polarization. UFAE and IAE Working Paper 766.09, 2009.

[12] The usual empirical finding is that the joint effect of stratification and clustering in samples design is to increase standard errors of the Gini index by considerable margin (see, e.g., [3]).

[18] Ezcurra R., Pascual P., Regional polarisation and national development in the European Union. *Urban Studies*, 44(1) (2007), pp. 99-122.
[19] Ezcurra R., Pascual P. and Rapún M., The dynamics of regional disparities in Central and Eastern Europe during transition. *European Planning Studies*, 15(10) (2007), pp. 1397-1421.
[20] Fedorov L., Regional inequality and regional polarization in Russia, 1990-99. *World Development*, 30(3) (2002), pp. 443-456.
[21] Gradín C., Polarization by sub-populations in Spain. *Review of Income and Wealth*, 46(4) (2000), pp. 457-474.
[22] Gradín C., Polarization and inequality in Spain: 1973-91. *Journal of Income Distribution*, 11(1) (2002), pp. 34-52.
[23] Gradín C. and Rossi M., Income distribution and income sources in Uruguay. *Journal of Applied Economics,* 9(1) (2006), pp. 49-69.
[24] Hussain M. A., The sensitivity of income polarization: time, length of accounting periods, equivalence scales, and income definitions, *Journal of Economic Inequality*, 7(3) (2008), pp. 207-223.
[25] Jenkins S. P., Did the middle class shrink during the 1980s? UK evidence from kernel density estimates. *Economics Letters*, 49(4) (1995), pp. 407-413.
[26] Kordos J., Lednicki B., Zyra M., The household sample surveys in Poland, *Statistics in Transition*, 5(4) (2002), pp. 555-589.
[27] Kot S. M., *Polaryzacja ekonomiczna. Teoria i zastosowanie*, Warszawa, PWN, 2008.
[28] Lasso de la Vega M. C. and Urrutia A. M., An alternative formulation of the Esteban-Gradín-Ray extended measure of polarization. *Journal of Income Distribution*, 15 (2006), pp. 42-54.
[29] Pagan A. and Ullah A., *Nonparametric Econometrics*, Cambridge University Press, New York, 1999.
[30] Rao C. R, *Linear Statistical Inference and Its Applications*, Second Edition, Wiley, New York, 1973.
[31] Ravallion M. and Chen S., What can new survey data tell us about recent changes in distribution and poverty? *World Bank Economic Review*, 11(2) (1997), pp. 357-382.
[32] Seshanna S. and Decornez S., Income polarization and inequality across countries: an empirical study. *Journal of Policy Modeling*, 25(4) (2003), pp. 335-358.
[33] Silverman B.W., *Density Estimatrion for Statistics and Data Analysis. Monographs on Statistics and Applied Probability*, Chapman and Hall, London, 1986.
[34] Wang Y.-Q. and Tsui K.-Y., Polarization orderings and new classes of polarization indices. *Journal of Public Economic Theory*, 2(3) (2000), pp. 349-363.
[35] Wolfson M. C., When inequalities diverge. *American Economic Review*, 84(2) (1994), pp. 353-358.
[36] Zhang X. and Kanbur R., What difference do polarisation measures make? An application to China, *Journal of Development Studies*, 37(3) (2001), pp. 85-98.

STATISTICAL INFERENCE ON CHANGES IN INCOME POLARIZATION IN POLAND

Summary

This paper estimates a popular measure of economic polarization (DER index) using Polish income micro-data from the Household Budget Survey (HBS) study for the period from 1998 to 2007. Using asymptotically distribution-free statistical inference we test whether the changes in the values of the estimated indices are statistically significant. Results show that during the period under study DER index has increased in the range from 4.9 to 6.5%, depending on the polarization-sensitivity parameter α. For the period 1998-2007, the changes in the DER index with $\alpha = 1$, for which polarization is empirically

distinguishable from inequality, are statistically significant at conventional levels. It is found, however, that for some sub-periods changes in the DER index are not statistically significant.

Keywords: polarization, income distribution, statistical inference, Poland

WNIOSKOWANIE STATYSTYCZNE O ZMIANACH POLARYZACJI DOCHODOWEJ W POLSCE

Streszczenie

W artykule dokonano estymacji popularnej miary polaryzacji ekonomicznej (indeks DER) dla danych o dochodach z polskich Badań Budżetów Gospodarstw Domowych w okresie 1998-2007. Używając niezależnych od rozkładu metod wnioskowania statystycznego przeprowadzono testy istotności różnic w wartościach estymowanych wskaźników. Wyniki pokazują, że w badanym okresie wartość indeksu DER wzrosła pomiędzy 4,9 a 6,5% w zależności od wartości parametru α mierzącego wagę przykładaną do polaryzacji. Zmiany indeksu DER dla $\alpha = 1$, dla którego polaryzacja zachowuje się empirycznie w sposób odmienny od nierówności, są w badanym okresie statystycznie istotne. W artykule pokazano jednak, że niektóre ze zmian indeksu DER dla krótszych okresów nie są istotne statystycznie.

PRZEGLĄD STATYSTYCZNY
R. LVIII – ZESZYT 1-2 – 2011

WITOLD ORZESZKO

WPŁYW REDUKCJI SZUMU LOSOWEGO METODĄ SCHREIBERA NA IDENTYFIKACJĘ SYSTEMU GENERUJĄCEGO DANE. ANALIZA SYMULACYJNA[1]

1. WSTĘP

Wszelkie rzeczywiste, w tym i ekonomiczne szeregi czasowe cechują się obecnością szumu losowego. Oczywiście fakt ten może znacząco utrudnić trafne identyfikowanie analizowanych zależności, co uzasadnia stosowanie metod redukcji szumu. Jednak z drugiej strony należy podkreślić, że wszelkie filtrowanie może prowadzić do zniszczenia bądź zniekształcenia zależności obecnych w analizowanych danych (por. Mees, Judd [17]).

Wiele spośród znanych metod identyfikacji nieliniowości i chaosu jest wrażliwych na obecność szumu losowego (por. np. Castagli i in. [3], Kantz, Schreiber [12], Shintani, Linton [22], Zeng i in. [25]). Z tego powodu metody redukcji szumu wykorzystywane są do filtrowania rzeczywistych danych jako wstępny etap w procesie identyfikacji chaosu lub nieliniowości. I choć nie jest to powszechnie stosowana praktyka, to otrzymane wyniki wydają się obiecujące (np. Harrison i in. [8], Hassani i in. [9, 10], Kantz i in. [11], Leontitsis i in. [14], Liang i in. [15], Orzeszko [18], Perc [19], Schreiber [21], Sivakumar i in. [23]).

Jedną z metod nieliniowej redukcji szumu losowego jest procedura zaproponowana przez Schreibera [20]. Przeprowadzone symulacje wykazały, że metoda ta, mimo swojej prostoty, skutecznie redukuje szum z szeregów wygenerowanych przez deterministyczne systemy o dynamice chaotycznej (Orzeszko [18], Schreiber [20], Sivakumar i in. [23]). Jednak należy podkreślić, że w przypadku analizy rzeczywistych szeregów czasowych badacz zwykle pozbawiony jest wiedzy o naturze systemu generującego. W szczególności, brak jest pewności, czy spełnione jest założenie leżące u podstaw metod redukcji szumu, tzn. o deterministycznym charakterze systemu generującego. Co więcej, to właśnie poznanie natury systemu generującego może być zasadniczym celem przeprowadzanego badania. Z tego powodu istotną kwestią jest stwierdzenie, jak dana metoda redukcji szumu zachowuje się w zastosowaniu do danych stochastycznych. Celem symulacji przeprowadzonych w niniejszej pracy było zweryfikowanie, czy prze-

[1] Badanie zostało sfinansowane przez Uniwersytet Mikołaja Kopernika w Toruniu w ramach grantu UMK nr 397-E „Identyfikacja nieliniowych zależności w ekonomicznych szeregach czasowych”.

filtrowanie szeregów losowych metodą Schreibera nie wprowadza do nich zależności, które mogłyby prowadzić do błędnych wniosków o naturze systemu generującego.

Praca skonstruowana jest następująco: w rozdziale drugim przedstawiono istotę i przebieg procesu redukcji szumu losowego, w rozdziałach trzecim oraz czwartym scharakteryzowano metody identyfikacji zależności nieliniowych w szeregach czasowych, tj. odpowiednio: test BDS oraz miarę informacji wzajemnej, w rozdziale piątym zaprezentowano przebieg oraz wyniki przeprowadzonych symulacji.

2. REDUKCJA SZUMU LOSOWEGO

Analizując określony rzeczywisty szereg czasowy (x_t) należy liczyć się z obecnością w nim szumu losowego, reprezentującego szum obserwacyjny, systemowy lub pewną ich kombinację. W takiej sytuacji w szeregu (x_t) można wyróżnić część deterministyczną (y_t) i stochastyczną (ε_t), co po zapisaniu w postaci addytywnej oznacza istnienie dekompozycji:

$$x_t = y_t + \varepsilon_t, \tag{1}$$

gdzie od (ε_t) wymaga się co najmniej posiadania szybko malejącej funkcji autokorelacji i nieskorelowania z sygnałem (y_t) (Kantz, Schreiber [12]).

Redukcja szumu losowego ma na celu upodobnienie „zanieczyszczonego" szeregu (x_t) do oryginalnego sygnału (y_t). Bardzo prostą, lecz mimo to stosunkowo skuteczną metodą redukcji szumu jest metoda Najbliższych Sąsiadów – NS (por. Schreiber [20]). W celu wyznaczenia wartości y_i (dla dowolnego i) metodą Najbliższych Sąsiadów należy skonstruować wektory opóźnień $\hat{x}_i = (x_{i-K},\ x_{i-K+1}, \ldots,\ x_i, \ldots,\ x_{i+L-1},\ x_{i+L})$, gdzie K i L są ustalonymi liczbami naturalnymi. Niech $\hat{x}_{l_1}, \hat{x}_{l_2}, \ldots, \hat{x}_{l_n}$ oznaczają n najbliższych (w sensie ustalonej $K + L + 1$ – wymiarowej metryki) sąsiadów wektora $\hat{x}_i$. W oparciu o wyznaczonych najbliższych sąsiadów, wartość $\tilde{y}_i$, będącą oszacowaniem y_i oblicza się ze wzoru:

$$\tilde{y}_i = \frac{1}{n} \sum_{i=1}^{n} x_{l_i}. \tag{2}$$

Dobór liczby najbliższych sąsiadów może odbywać się wprost (jak to zostało zaprezentowane w przedstawionym powyżej algorytmie) lub przez zadanie promienia sąsiedztwa r.

Metoda NS zazwyczaj prowadzi do różnych szeregów w zależności od przyjętych wartości parametrów K, L oraz r(lub n). Do wyboru właściwych wartości parametrów, a co za tym idzie – do wyboru odpowiedniego szeregu wynikowego, można zastosować wskaźnik poziomu redukcji szumu NRL (*Noise Reduction Level*) (Orzeszko [18]). Metoda ta opiera się na obserwowanej zależności pomiędzy strukturą geometryczną atraktora a siłą szumu dodawanego do systemu. Zależność ta polega na „pogrubianiu" atraktora, tzn., średnio rzecz biorąc, oddalaniu się bliskich sobie stanów wraz ze wzrostem siły szumu.

Punktem wyjścia metody jest rekonstrukcja atraktora systemu w oparciu o szereg (x_t), polegająca na skonstruowaniu m-wymiarowych wektorów opóźnień, gdzie m jest ustaloną liczbą naturalną. Niech T oznacza liczbę wszystkich dostępnych wektorów opóźnień, natomiast d_i $(i = 1, 2, ..., T)$ – odległość i–tego oczyszczonego wektora opóźnień $(\tilde{y}_{i-K}, \tilde{y}_{i-K+1}, ..., \tilde{y}_i, ..., \tilde{y}_{i+L-1}, \tilde{y}_{i+L})$ do jego najbliższego sąsiada. Wówczas wielkość $d_{\min}$, określoną wzorem $d_{\min} = \frac{1}{T}\sum_{i=1}^{T} d_i$, można interpretować jako miarę grubości zrekonstruowanego atraktora systemu. Naczelną ideą proponowanej metody jest wskazanie spośród szeregów, będących wynikiem redukcji szumu, takiego, dla którego względny spadek wartości $d_{\min}$ (w stosunku do zanieczyszczonego szeregu) jest największy. Należy jednak zauważyć, że uwzględnienie wyłącznie tego czynnika byłoby niewystarczające, gdyż prowadziłoby zawsze do wyboru szeregu złożonego z równych (lub prawie równych) wartości[2]. Z tego powodu do wskaźnika NRL dodatkowo wprowadzono wartość $\left|\frac{diam - diam^0}{diam^0}\right|$, gdzie $diam^0$ i $diam$ oznaczają średnice zrekonstruowanego atraktora, odpowiednio, przed i po redukcji szumu. Rolą względnej zmiany średnicy jest zmierzenie poziomu deformacji oczyszczonego atraktora, spowodowanej przez redukcję szumu.

Ostatecznie wskaźnik poziomu redukcji szumu NRL wyraża się wzorem:

$$NRL = \frac{d_{\min} - d^0_{\min}}{d^0_{\min}} + \left|\frac{diam - diam^0}{diam^0}\right|. \tag{3}$$

Spośród szeregów otrzymanych w wyniku redukcji szumu należy wybrać taki, który cechuje się najniższym wskaźnikiem NRL.

Z samej konstrukcji wartość NRL zależy od przyjętego wymiaru zanurzenia m. Zgodnie z kryterium Takensa [1981], w procesie rekonstrukcji przestrzeni stanów należy uwzględnić wektory opóźnień, których wymiar spełnia warunek $m > 2d$, gdzie d jest wymiarem przestrzeni stanów.

3. TEST BDS

BDS (Brock i in. [1]) jest nieparametrycznym testem weryfikującym hipotezę zerową, że badany szereg czasowy jest realizacją procesu i.i.d. (tzn. procesu niezależnych zmiennych losowych o jednakowym rozkładzie). W swojej konstrukcji odwołuje się do całki korelacyjnej, zadanej wzorem:

$$C^T_m(\varepsilon) = \lim_{T\to\infty} \frac{\left|\left\{(i,j);\ \ \left\|\hat{x}_i - \hat{x}_j\right\| < \varepsilon,\ 1 \leqslant i, j \leqslant T\right\}\right|}{T^2}, \tag{4}$$

[2] W przypadku metody NS szereg taki otrzymuje się dla odpowiednio dużej wartości parametru r (lub n).

gdzie symbol || oznacza moc zbioru, natomiast T jest liczbą wszystkich m-wymiarowych wektorów opóźnień.

Teoretyczną podstawą testu BDS jest asymptotyczna zależność pomiędzy całkami korelacyjnymi $C_m^T(\varepsilon)$ i $C_1^T(\varepsilon)$. Udowodniono, że dla każdego $m > 1$ oraz $\varepsilon > 0$ wartość $C_m^T(\varepsilon)$ szeregu i.i.d. zbiega do $\left(C_1^T(\varepsilon)\right)^m$ przy $T \to \infty$. Mająca tu zastosowanie statystyka W określona jest wzorem[3]:

$$W_{T,m}(\varepsilon) = \frac{\sqrt{T}\left(C_m^T(\varepsilon) - \left(C_1^T(\varepsilon)\right)^m\right)}{\sigma_{T,m}(\varepsilon)}. \tag{5}$$

Przy założeniu prawdziwości hipotezy zerowej, że szereg jest realizacją procesu i.i.d., statystyka W jest zbieżna według rozkładu do N(0, 1) przy $T \to \infty$ (por. Kanzler [13]). BDS jest testem dwustronnym, w którym weryfikacji podlega hipoteza zerowa o nieistotności statystyki W.

Należy podkreślić, że test BDS identyfikuje zależności zarówno liniowe, jak i nieliniowe, więc w celu identyfikacji nieliniowości należy badać szereg uprzednio przefiltrowany odpowiednim modelem ARMA. Zaletą tego testu jest fakt, że identyfikuje nieliniowości bardzo różnej natury.

Test BDS jest wrażliwy na liczbę obserwacji. Brock i in. [2] w oparciu o przeprowadzone symulacje stwierdzili, że wiarygodne wyniki otrzymuje się dla szeregów liczących co najmniej 250 obserwacji. Dodatkowo, w przypadku krótkich szeregów (do 500 obserwacji) w celu wyznaczania wartości krytycznych bezpieczniej jest stosować procedurę *bootstrap* niż tablice rozkładu normalnego (por. Brock i in. [2], Kanzler [13]).

W oparciu o dokonane symulacje Brock i in. [2] proponują, aby parametr m przyjmował w teście wartości 2, 3, 4, 5 dla krótkich szeregów (do 500 obserwacji) i wartości 2, 3, ..., 10 – dla długich (co najmniej 2000 obserwacji). Ponadto, z symulacji tych wynika postulat, aby za ε przyjmować wartości pomiędzy $0,5 \cdot \sigma$ i $1,5 \cdot \sigma$, gdzie σ jest odchyleniem standardowym badanego szeregu. Z kolei z badań przeprowadzonych przez L. Kanzlera [13] wynika, że przyjmowana w teście BDS wartość parametru ε powinna zależeć od typu rozkładu analizowanego szeregu. Przykładowo w przypadku szeregów o rozkładzie normalnym lub do niego zbliżonego, należy przyjąć $\varepsilon = 1,5 \cdot \sigma$ lub $\varepsilon = 2 \cdot \sigma$.

4. MIARA INFORMACJI WZAJEMNEJ

Miara informacji wzajemnej (ang. Mutual Information – MI) jest jedną z najważniejszych metod pomiaru nieliniowych zależności w szeregach czasowych. Określona jest ona wzorem:

[3] $\sigma_{T,m}(\varepsilon)$ jest czynnikiem normalizującym, którego wzór przedstawiony jest np. w Brock i in. [1991].

$$I(X,Y) = \iint p(x,y) \log\left(\frac{p(x,y)}{p_1(x)p_2(y)}\right) dxdy, \tag{6}$$

gdzie $p(x,y)$ jest funkcją gęstości rozkładu łącznego, natomiast $p_1(x)$ oraz $p_2(y)$ są gęstościami brzegowymi zmiennych X i Y.

Można udowodnić, że $I(X,Y)$ przyjmuje zawsze wartości nieujemne oraz $I(X,Y) = 0$ wtedy i tylko wtedy, gdy X i Y są niezależne (np. Granger, Lin [7]). Oznacza to, że test niezależności bazujący na mierze informacji wzajemnej polega na weryfikacji hipotezy H_0: $I(X,Y) = 0$ przeciw H_1: $I(X,Y) > 0$.

W literaturze przedmiotu istnieją różne propozycje szacowania wartości $I(X,Y)$. Zasadniczo, sprowadzają się one do oszacowania funkcji gęstości $p(x,y)$, $p_1(x)$ oraz $p_2(y)$ (por. wzór 6). Proponowane metody estymacji można podzielić na trzy główne grupy (por. Dionisio, Menezes, Mendes [4]):

- odwołujące się do histogramu,
- oparte na estymatorach jądrowych,
- metody parametryczne.

Miara MI identyfikuje zależności zarówno liniowe, jak i nieliniowe, więc podobnie jak w przypadku testu BDS, aby zidentyfikować zależności nieliniowe, badany szereg trzeba przefiltrować modelem typu ARMA. Miarę tę można wykorzystać również do pomiaru autozależności w pojedynczym szeregu czasowym (x_t). W tym celu za (y_t) należy przyjąć szereg opóźnionych wartości (x_t).

5. ANALIZA SYMULACYJNA

5.1 PRZEBIEG BADANIA

Celem przeprowadzonych symulacji było zweryfikowanie, czy dokonanie redukcji szumu losowego metodą Najbliższych Sąsiadów może zakłócić identyfikację zależności nieliniowych w szeregach czasowych, a dokładniej mówiąc: czy metoda ta może wprowadzić do szeregów losowych zależności o charakterze nieliniowym.

Badania oparto na wygenerowanych w programie Matlab 6.5 szeregach liczb pseudolosowych o rozkładzie N(0, 1): 1000 szeregów krótkich (300 obserwacji) oraz 1000 szeregów długich (2000 obserwacji).

W pierwszej kolejności, każdy z wygenerowanych szeregów poddano redukcji szumu losowego przy zastosowaniu metody NS. Następnie przefiltrowane szeregi przetestowano pod kątem ich niezależności. W tym celu zastosowano procedurę opartą na teście BDS, miarę informacji wzajemnej oraz współczynnik korelacji liniowej Pearsona.

5.2 REDUKCJA SZUMU LOSOWEGO

W redukcji szumu losowego zastosowano metodę Najbliższych Sąsiadów.[4] W badaniu arbitralnie przyjęto wartości parametrów $K = 2$ oraz $L = 2$, natomiast promień otoczenia r wyznaczono przy użyciu wskaźnika NRL.[5]

Przy wyznaczaniu wskaźnika NRL analizie poddano po jednym z wygenerowanych długich i krótkich szeregów. W badaniu rozważono osiem wartości promienia otoczenia: r = 0,2; 0,3; 0,4; 0,5; 0,7; 1; 1,5; 2. W celu wyznaczenia optymalnego promienia obliczono wartości wskaźnika NRL(m), kolejno, dla m = 1, 2, 3, 5, 7, 10. Obliczone wartości NRL zaprezentowano w Tabelach 1-2. Dla każdego m pogrubioną czcionką zaznaczono najmniejszą wartość wskaźnika NRL(m). Dodatkowo w każdej komórce obydwu tych tabel przedstawiono w nawiasach wartości drugiej składowej wskaźnika NRL, tzn. względnej zmiany średnicy – $\Delta diam$ [%], umożliwiające ocenę stopnia deformacji przefiltrowanego szeregu.

Tabela 1

Wartości NRL dla szeregu losowego N(0,1), złożonego z 300 obserwacji

NRL / r	NRL(1)	NRL(2)	NRL(3)	NRL(5)	NRL(7)	NRL(10)
0,2	0,0000 /0,00/	0,3642 /0,00/	0,0417 /0,00/	0,1551 /0,00/	-0,0021 /0,00/	-0,0351 /0,00/
0,3	0,0000 /0,00/	1,6894 /0,00/	0,0509 /0,00/	1,0959 /0,00/	0,2353 /0,00/	0,2390 /0,00/
0,4	0,0319 /0,00/	-0,0749 /0,00/	0,9184 /0,00/	0,7479 /0,00/	1,1614 /0,00/	0,5870 /0,00/
0,5	0,0427 /0,00/	-3,8614 /0,00/	-0,3144 /0,00/	0,5066 /0,00/	1,5307 /0,02/	2,8491 /0,43/
0,7	3,9391 /0,00/	-8,2530 /0,00/	-9,0608 /0,00/	-10,7853 /0,00/	-7,2211 /0,39/	-1,1949 /5,07/
1	6,7304 /4,16/	**-14,9189** /3,28/	**-13,9593** /7,46/	**-14,8774** /10,88/	**-16,0313** /10,55/	**-9,8927** /13,39/
1,5	-1,6249 /38,82/	-7,3907 /39,18/	-7,9031 /41,91/	-10,7387 /39,20/	-10,9918 /39,51/	-5,1402 /44,98/
2	**-5,5032** /52,41/	-12,5752 /52,74/	-12,9643 /55,27/	-9,6012 /57,76/	-10,1073 /57,31/	-5,8164 /61,35/

Źródło: obliczenia własne

Biorąc pod uwagę otrzymane rezultaty, do metody redukcji szumu losowego przyjęto wartość $r = 1$ dla szeregów krótkich oraz $r = 0,7$ dla szeregów długich. W przypadku szeregów krótkich, wybór $r = 1$ nie pozostawiał wątpliwości; jedynie dla $m = 1$ ten poziom parametru r nie prowadził do najmniejszej wartości wskaźnika NRL(m). Nieco mniej jednoznaczne wskazanie otrzymano dla szeregów długich. Jed-

[4] Obliczenia wykonano w programie Matlab 6.5 w oparciu o procedurę napisaną przez A. Leontitsisa.
[5] Obliczenia wykonano w programie Excel w oparciu o procedurę własną.

Tabela 2

Wartości NRL dla szeregu losowego N(0,1), złożonego z 2000 obserwacji

NRL / r	NRL(1)	NRL(2)	NRL(3)	NRL(5)	NRL(7)	NRL(10)
0,2	0,0340 /0,00/	-0,2248 /0,00/	-0,2762 /0,00/	-0,1610 /0,00/	0,0440 /0,00/	0,0965 /0,00/
0,3	-1,2391 /0,00/	-0,5434 /0,00/	-1,7409 /0,00/	-1,0148 /0,00/	-0,0353 /0,00/	0,4910 /0,45/
0,4	1,1279 /0,00/	0,7097 /0,00/	-2,5220 /0,00/	-5,4955 /0,00/	-1,1727 /0,88/	-1,1857 /0,19/
0,5	-1,6668 /0,00/	-1,9316 /0,00/	-4,6991 /0,84/	-9,3423 /0,00/	-3,9642 /1,67/	-3,6049 /1,70/
0,7	8,0467 /4,19/	-7,4150 /0,00/	**-12,3525** /0,84/	**-13,1207** /4,94/	**-13,3316** /2,18/	-8,0552 /5,85/
1	1,9431 /10,79/	-5,5266 /13,37/	-11,2943 /14,95/	-8,4624 /19,54/	-7,0748 /20,89/	-7,4781 /20,17/
1,5	**-3,4391** /35,26/	-6,1141 /36,96/	-7,0012 /39,91/	-6,2428 /41,95/	-5,0225 /43,22/	-6,9233 /41,44/
2	-2,1241 /50,58/	**-9,6498** /52,79/	-10,4288 /55,29/	-9,6989 /57,58/	-8,1550 /59,32/	**-9,3679** /58,31/

Źródło: obliczenia własne

nak i w tej sytuacji zauważalna jest wyższość parametru $r = 0,7$ nad pozostałymi analizowanymi poziomami.

Metoda NS pozostawia nieprzefiltrowane K pierwszych oraz L ostatnich obserwacji, co skutkuje tym, że szeregi krótkie zostały skrócone do 296 a długie – do 1996 obserwacji.

5.3 TEST BDS

Dla każdego z przefiltrowanych szeregów wyznaczono statystykę W, przyjmując kolejno wartości wymiaru zanurzenia $m = 2, 3, 4, 5$.[6] Wartość ε ustalono na poziomie $\varepsilon = 1,5\sigma$, gdzie σ jest odchyleniem standardowym analizowanego szeregu.

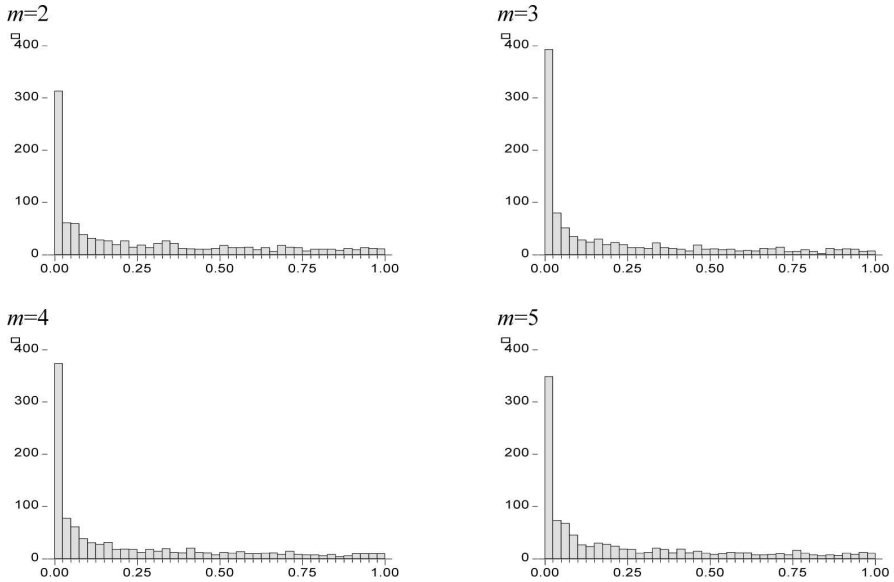
Dla poszczególnych analizowanych szeregów, w odniesieniu do kolejnych rozważanych wartości parametru m, wyznaczono wartość empirycznego poziomu istotności (ang. p-value), umożliwiającą zweryfikowanie hipotezy zerowej o nieistotności statystyki W (tzn. o braku zależności w szeregu). Wartość ta została wyznaczona w oparciu o procedurę bootstrap (1000 iteracji).[7]

Na Rysunkach 1-2 przedstawiono histogramy empirycznych poziomów istotności obliczonych dla różnych wartości parametru m. Ponadto, w oparciu o otrzymane hi-

[6] Obliczenia wykonano w programie Matlab 6.5 w oparciu o procedurę napisaną przez L. Kanzlera.

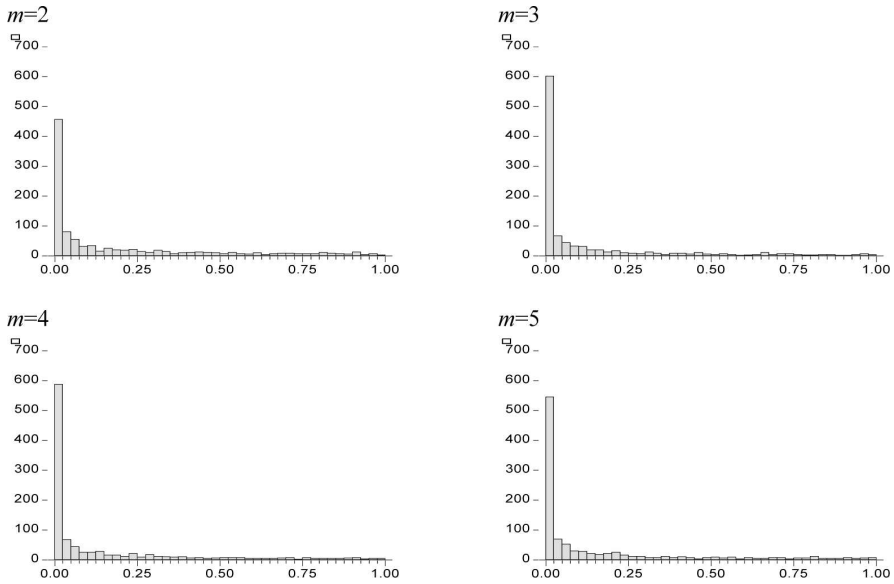
[7] Zastosowano procedurę bootstrap bez powtarzania. Wyznaczone empiryczne poziomy istotności odpowiadają testowi dwustronnemu (tzn. są dwukrotnością grubości odpowiedniego ogona rozkładu empirycznego).

stogramy wyznaczono odsetek szeregów, dla których odrzucono hipotezę zerową, przy wybranych poziomach istotności $\alpha = 1\%,\ 5\%,\ 10\%$. Wyniki te zostały przedstawione w Tabelach 3-4.



Rysunek 1. Histogramy empirycznych poziomów istotności statystyki W dla krótkich szeregów

Źródło: obliczenia własne



Rysunek 2. Histogramy empirycznych poziomów istotności statystyki W dla długich szeregów

Źródło: obliczenia własne

Tabela 3

Odsetek krótkich szeregów, dla których odrzucono H_0

	m=2	m=3	m=4	m=5
α = 1%	23,6%	31,4%	29,5%	26,7%
α = 5%	38,0%	47,4%	45,7%	42,8%
α = 10%	47,2%	56,0%	55,4%	53,6%

Źródło: obliczenia własne

Tabela 4

Odsetek długich szeregów, dla których odrzucono H_0

	m=2	m=3	m=4	m=5
α = 1%	37,8%	52,8%	50,6%	46,9%
α = 5%	54,0%	67,1%	65,8%	61,6%
α = 10%	62,2%	74,4%	72,4%	69,4%

Źródło: obliczenia własne

W otrzymanych rozkładach empirycznych zdecydowanie dominują wartości bliskie zeru. Potwierdzają to wyniki zaprezentowane na Rysunkach 1-2 oraz w Tabelach 3-4. Jak widać na wykresach, w przypadku szeregów krótkich – ponad 30%, a w przypadku szeregów długich – nawet ponad 45% obliczonych empirycznych poziomów istotności nie przekracza wartości 0,025. Z kolei obliczenia zestawione w Tabelach 3-4 pokazują, że odsetek szeregów, dla których odrzucono hipotezę zerową znacznie przewyższa przyjęty krytyczny poziom istotności α. Z otrzymanych rezultatów testu BDS wynika więc, że przefiltrowanie szeregów metodą NS powoduje pojawienie się w wielu z nich autozależności pomiędzy obserwacjami. Naturalnie na tym etapie badania nie można wyciągnąć wniosków dotyczących charakteru tych zależności.[8]

5.4 MIARA INFORMACJI WZAJEMNEJ

Dla każdego z przefiltrowanych szeregów oszacowano wartości miary informacji wzajemnej dla opóźnień czasowych k = 1, 2, ..., 10.[9] Do oszacowania miary MI wykorzystano metodę zaproponowaną przez Frasera, Swinneya [5]. Jest to metoda oparta na analizie dwuwymiarowego histogramu. Mówiąc w uproszczeniu, polega ona na podziale przestrzeni dwuwymiarowej zawierającej pary (x_t, y_t) na prostokątne komórki (według określonej zasady) i obliczeniu, jaki odsetek naniesionych punktów znajduje się w każdej z komórek. Następnie znajduje zastosowanie wzór (6), przy czym ob-

[8] Zależności te mogą być zarówno liniowe, jak i nieliniowe. Kwestii tej został poświęcony podrozdział 5.5.

[9] Obliczenia wykonano w programie Matlab 6.5 w oparciu o procedurę napisaną przez A. Leontitsisa.

liczone odsetki są oszacowaniem funkcji gęstości, natomiast całkowanie odbywa się numerycznie.

Tak jak w przypadku testu BDS, w celu zweryfikowania hipotezy o nieistotności miary MI (tzn. o braku zależności w szeregu) obliczono empiryczne poziomy istotności. Wartości te zostały wyznaczone w oparciu o procedurę bootstrap (10 000 iteracji).[10] Dla każdego z szeregów otrzymano w ten sposób 10 000 wartości MI(1), których rozkład był podstawą do wyznaczenia empirycznych poziomów istotności dla każdego kolejnego $k = 1, 2, ..., 10$.

Na Rysunkach 3-4 przedstawiono histogramy empirycznych poziomów istotności obliczonych dla różnych wartości parametru k. Ponadto, w oparciu o otrzymane histogramy wyznaczono odsetek szeregów, dla których odrzucono hipotezę zerową, przy wybranych poziomach istotności $\alpha = 1\%,\ 5\%,\ 10\%$. Wyniki te zostały przedstawione w Tabelach 5-6.

Tabela 5

Odsetek krótkich szeregów, dla których odrzucono H_0

	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
α = 1%	44,8%	41,3%	1,1%	1,3%	1,1%	1,0%	1,8%	1,5%	1,6%	1,8%
α = 5%	66,5%	61,6%	6,3%	5,7%	6,1%	6,1%	5,5%	7,5%	6,0%	6,8%
α = 10%	76,4%	72,1%	12,1%	11,6%	11,2%	10,6%	11,2%	13,6%	10,8%	14,5%

Źródło: obliczenia własne

Tabela 6

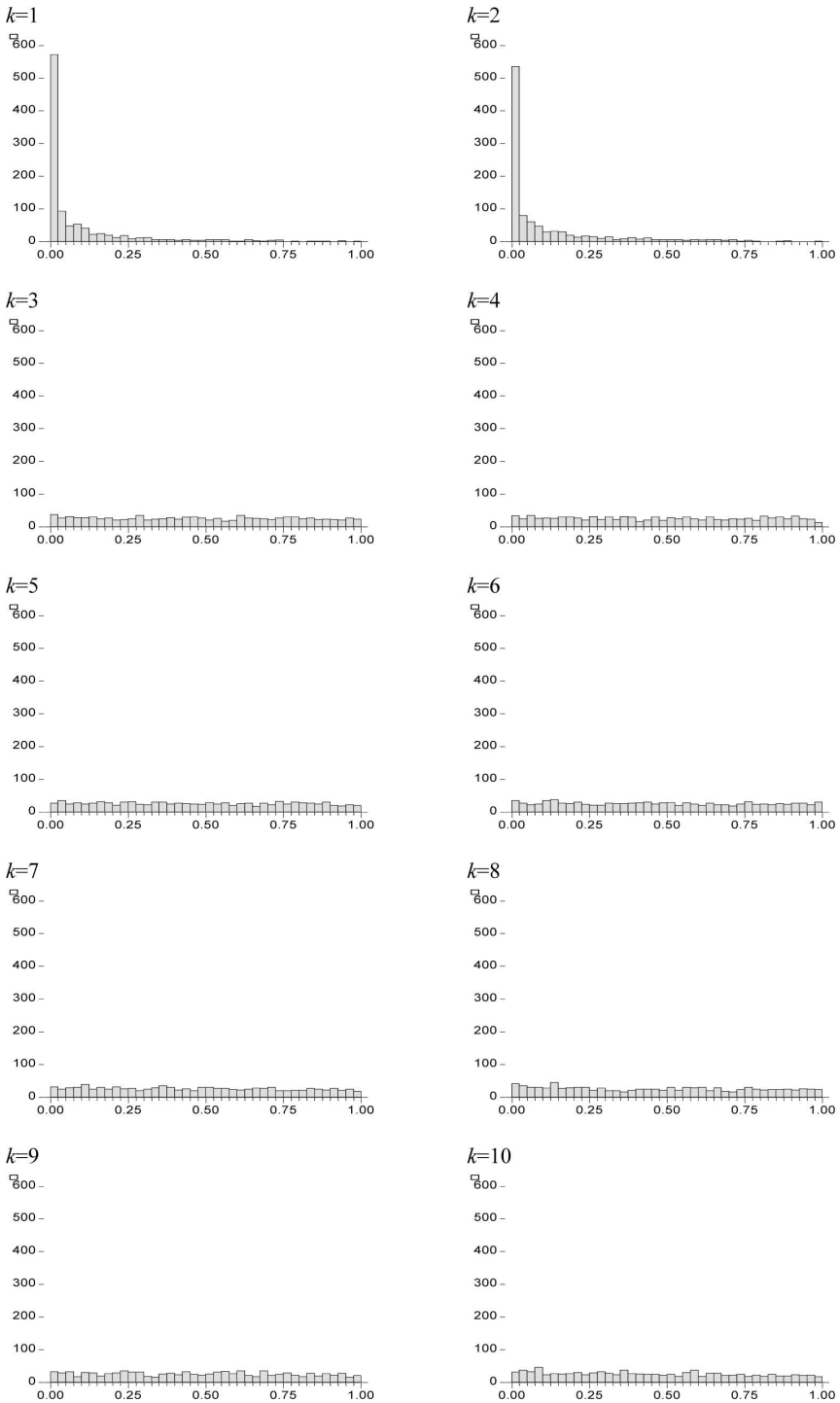
Odsetek długich szeregów, dla których odrzucono H_0

	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
α = 1%	73,7%	66,1%	0,8%	0,8%	0,8%	0,7%	1,1%	1,0%	1,3%	1,3%
α = 5%	88,6%	81,4%	5,2%	5,3%	4,5%	5,0%	4,9%	5,9%	6,1%	6,1%
α = 10%	93,7%	89,8%	9,7%	9,2%	9,1%	10,3%	10,3%	10,9%	10,2%	11,6%

Źródło: obliczenia własne

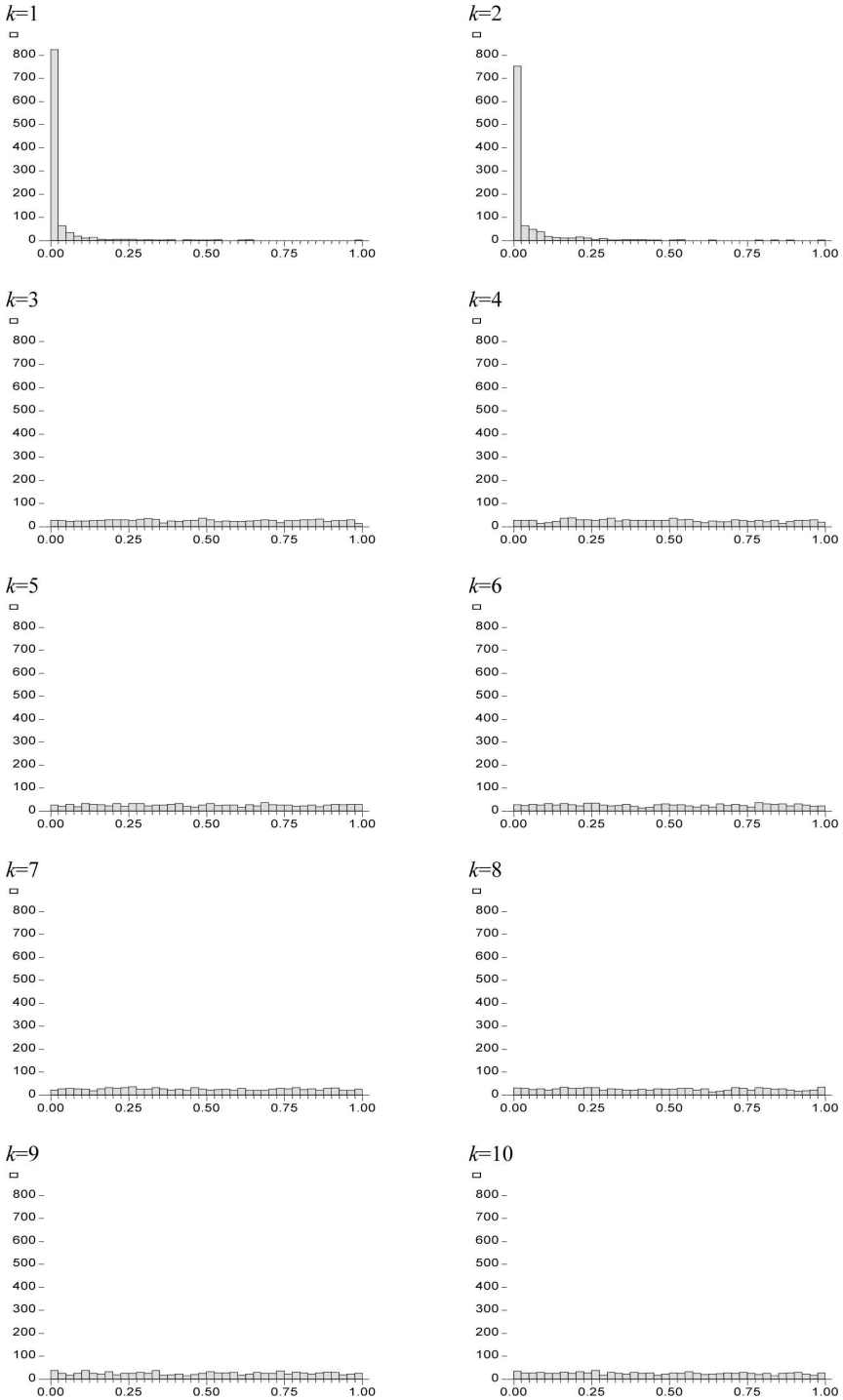
Z Rysunków 3-4 dla $k = 1$ oraz $k = 2$ wynika, że w przypadku szeregów krótkich – ponad 50%, a w przypadku szeregów długich – nawet ponad 75% wyznaczonych empirycznych poziomów istotności nie przekracza wartości 0,025. Istnienie autozależności potwierdzają również obliczenia zaprezentowane w Tabelach 5-6. Pokazują one, że dla $k = 1$ oraz $k = 2$ odsetek szeregów, dla których odrzucono hipotezę zerową znacznie przewyższa przyjęty krytyczny poziom istotności α. Z kolei w przypadku wyższych opóźnień (tzn. $k > 2$), histogramy są płaskie, a widoczne w Tabelach 5-6

[10] Zastosowano procedurę bootstrap bez powtarzania. Wyznaczone empiryczne poziomy istotności odpowiadają testowi prawostronnemu (tzn. są grubością prawego ogona rozkładu empirycznego).



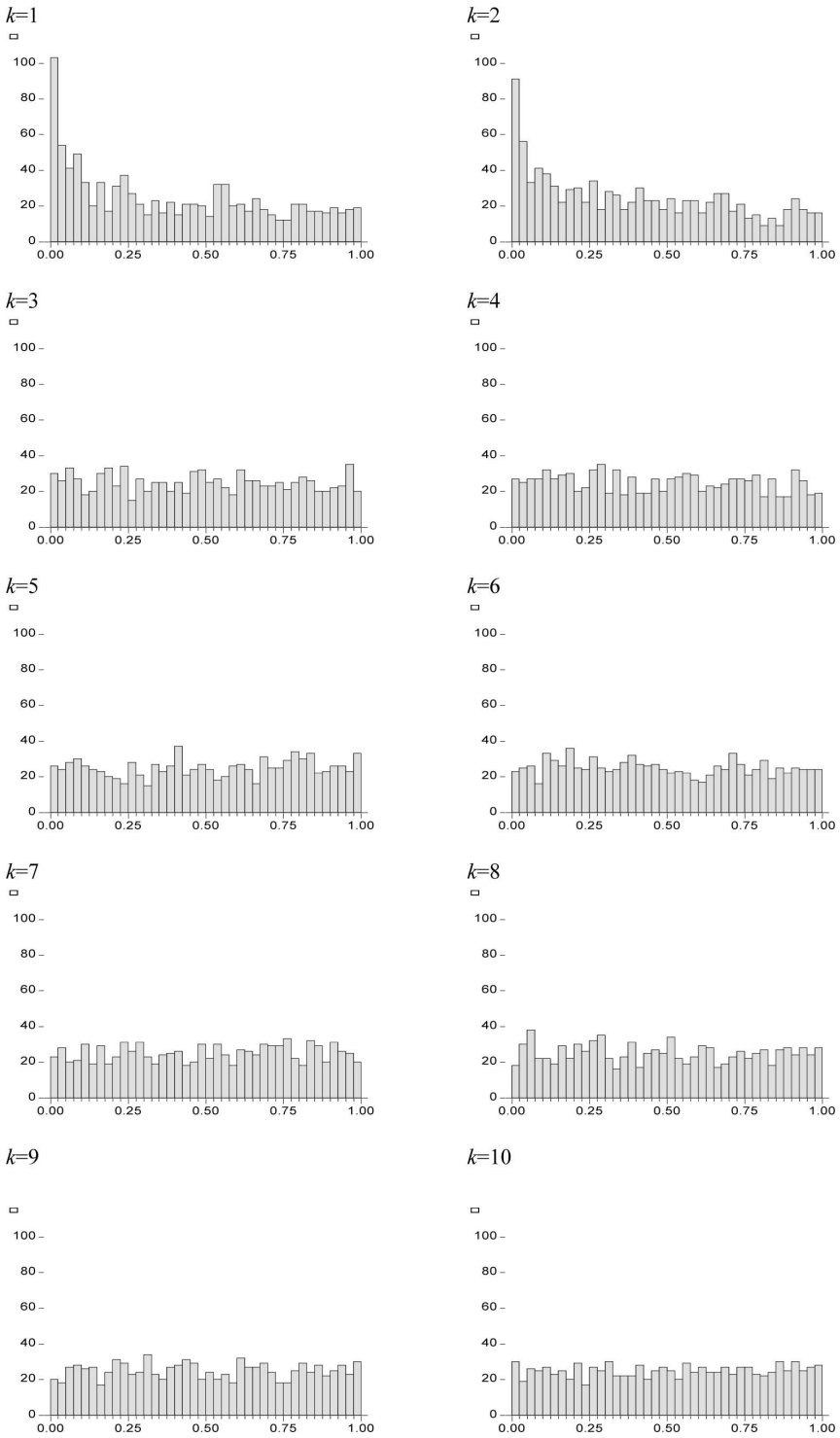
Rysunek 3. Histogramy empirycznych poziomów istotności miary MI dla krótkich szeregów

Źródło: obliczenia własne



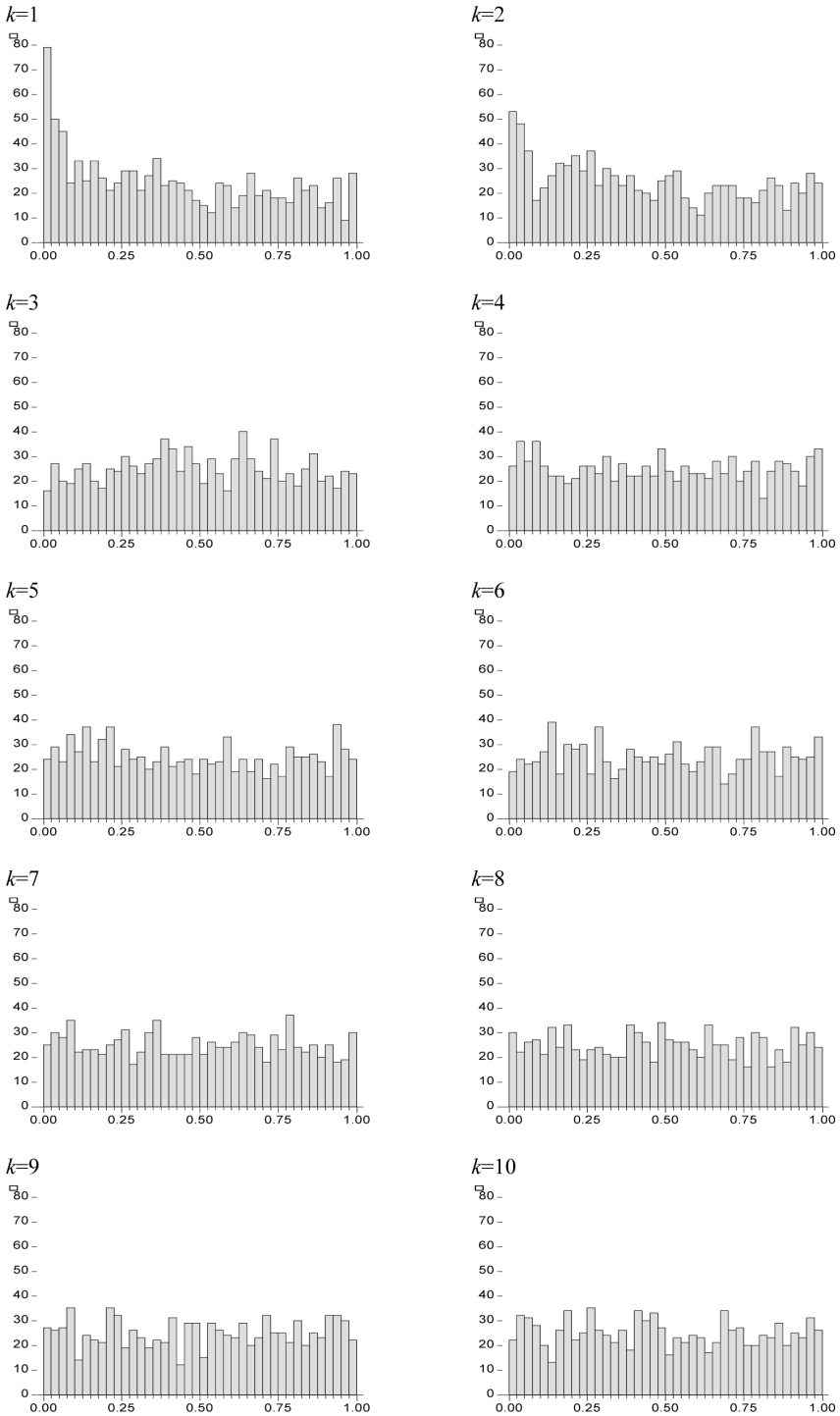
Rysunek 4. Histogramy empirycznych poziomów istotności miary MI dla długich szeregów

Źródło: obliczenia własne



Rysunek 5. Histogramy empirycznych poziomów istotności współczynnika AC dla krótkich szeregów

Źródło: obliczenia własne



Rysunek 6. Histogramy empirycznych poziomów istotności współczynnika AC dla długich szeregów

Źródło: obliczenia własne

odsetki szeregów, dla których odrzucono hipotezę zerową, są zbliżone do przyjętych poziomów istotności. Otrzymane rezultaty wyraźnie wskazują, że w analizowanych szeregach pojawiły się autozależności pierwszego oraz drugiego rzędu. Istnienie w szeregach zależności do drugiego rzędu włącznie wydaje się być konsekwencją ustalenia wartości parametrów $K = 2$ i $L = 2$ w metodzie NS.

5.5 ANALIZA LINIOWOŚCI

Zidentyfikowane przez test BDS oraz miarę MI zależności w przefiltrowanych szeregach mogą być różnej natury. Celem niniejszego podrozdziału jest odpowiedź na pytanie, czy mają one charakter nieliniowy.

W tym celu dla każdego z analizowanych szeregów obliczono współczynnik autokorelacji liniowej Pearsona, dla opóźnień czasowych $k = 1, 2, ..., 10$.[11]

Na Rysunkach 5-6 przedstawiono histogramy empirycznych poziomów istotności obliczonych dla różnych wartości parametru k. Ponadto, w oparciu o otrzymane histogramy wyznaczono odsetek szeregów, dla których odrzucono hipotezę zerową o nieistotności współczynnika, przy wybranych poziomach istotności $\alpha = 1\%,\ 5\%,\ 10\%$. Wyniki te zostały przedstawione w Tabelach 7-8.

Tabela 7

Odsetek krótkich szeregów, dla których odrzucono H_0

	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
α = 1%	5,9%	5,5%	1,3%	1,0%	1,1%	1,0%	1,0%	0,7%	0,6%	1,2%
α = 5%	15,7%	14,7%	5,6%	5,2%	5,0%	4,8%	5,1%	4,8%	3,8%	4,9%
α = 10%	24,7%	22,1%	11,6%	10,6%	10,8%	9,0%	9,2%	10,8%	9,3%	10,0%

Źródło: obliczenia własne

Tabela 8

Odsetek długich szeregów, dla których odrzucono H_0

	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
α = 1%	4,3%	2,9%	0,6%	0,6%	0,5%	1,0%	0,8%	1,5%	0,7%	0,8%
α = 5%	12,9%	10,1%	4,3%	6,2%	5,3%	4,3%	5,5%	5,2%	5,3%	5,4%
α = 10%	19,8%	15,5%	8,2%	12,6%	11,0%	8,8%	11,8%	10,5%	11,5%	11,3%

Źródło: obliczenia własne

Podobnie jak w przypadku miary MI otrzymane rezultaty wskazują na obecność autozależności pierwszego i drugiego rzędu. Jednak w tym wypadku, odsetek szeregów ze zidentyfikowanymi zależnościami jest dla każdego zaprezentowanego poziomu

[11] Do tego celu wykorzystano funkcję corrcoef w programie Matlab 6.5.

istotności dużo niższy niż w przypadku metody MI. Prowadzi to do wniosku, że tylko niewielką część zidentyfikowanych przez MI zależności można tłumaczyć obecnością autokorelacji liniowej. Oznacza to, że w przypadku większości przefiltrowanych szeregów zidentyfikowane autozależności mają charakter nieliniowy.

5.6 PODSUMOWANIE WYNIKÓW SYMULACJI

Przeprowadzone symulacje wskazały, że szeregi losowe po dokonaniu redukcji szumu metodą NS mogą wykazywać obecność zależności nieliniowych. Naturalnie wniosek ten zasadniczo podważa zasadność stosowania analizowanej metody w procesie identyfikacji nieliniowości w rzeczywistych szeregach czasowych. Oznacza on bowiem, że zidentyfikowane zależności mogą nie pochodzić z systemu generującego lecz być efektem filtrowania danych. Z otrzymanego wniosku wynika, że w procesie analizy rzeczywistych szeregów czasowych metodę Schreibera należy stosować z dużą ostrożnością. Przede wszystkim nie należy stosować redukcji szumu, gdy brak jest wyraźnych podstaw do stwierdzenia, że analizowane dane pochodzą z systemu deterministycznego. Z otrzymanych wyników płynie też wniosek ogólniejszej natury – o potrzebie przebadania innych metod redukcji szumu pod kątem ich zachowania w zastosowaniu do szeregów stochastycznych.

Uniwersytet Mikołaja Kopernika

LITERATURA

[1] Brock W.A., Dechert W.D., Scheinkman J.A. [1987], A test for independence based on the correlation dimension, SSRI Working Paper no. 8702, Department of Economics, University of Wisconsin, Madison.

[2] Brock W.A., Hsieh D.A., LeBaron B. [1991], *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*, The MIT Press, Cambridge, Massachusetts, London.

[3] Castagli M., Eubank S., Farmer J.D., Gibson J. [1991], State space reconstruction in the presence of noise, *Physica D*, 51, 52-98.

[4] Dionisio A., Menezes R., Mendes D.A. [2003], Mutual Information: a dependence measure for nonlinear time series, working paper.

[5] Fraser A.M., Swinney H.L. [1986], Independent coordinates for strange attractors from mutual information, *Physical Review A*, 33.2, 1134-1140.

[6] Granger C. W. J., Terasvirta T. [1993], *Modelling nonlinear economic relationship*, Oxford University Press.

[7] Granger C. W. J., Lin J-L. [1994], Using the mutual information coefficient to identify lags in nonlinear models, *Journal of Time Series Analysis*, 15, 371-384.

[8] Harrison R.G., Yua D., Oxleyb L., Lua W., Georgec D. [1999], Non-linear noise reduction and detecting chaos: some evidence from the S&P Composite Price Index, *Mathematics and Computers in Simulation*, 48, 497-502.

[9] Hassani H., Dionisio A., Ghodsi M. [2009a], The effect of noise reduction in measuring the linear and nonlinear dependency of financial markets, *Nonlinear Analysis: Real World Applications*, doi:10.1016/j.nonrwa.2009.01.004.

[10] Hassani H., Zokaeib M., Rosenc D., Amiric S., Ghodsi M. [2009b], Does noise reduction matter for curve fitting in growth curve models?, *Computer methods and programs in biomedicine*, 96, 173-181.
[11] Kantz H., Schreiber T., Hoffman I. [1993], Nonlinear noise reduction: A case study on experimental data, *Physical Review E*, 48.2, 1529-1538.
[12] Kantz H., Schreiber T. [1997], *Nonlinear time series analysis*, Cambridge University Press
[13] Kanzler L. [1999], Very Fast and Correctly Sized Estimation of the BDS Statistic, Department of Economics, Oxford University.
[14] Leontitsis A., Bountis T., Pagge J. [2004], An adaptive way for improving noise reduction using local geometric projection, *Chaos* 14, 106-110.
[15] Liang H., Lin Z., Yin F., [2005], Removal of ECG contamination from diaphragmatic EMG by nonlinear filtering, *Nonlinear Analysis*, 63, 745-753.
[16] Maasoumi E., Racine J. [2002], Entropy and predictability of stock market returns, *Journal of Econometrics*, 107, 291-312.
[17] Mees A. I., Judd K. [1993], Dangers of geometric filtering, *Physica D*, 68, 427-436.
[18] Orzeszko W. [2008], The New Method of Measuring the Effects of Noise Reduction in Chaotic Data, *Chaos Solitons and Fractals*, 38, 1355-1368.
[19] Perc M., [2005], Nonlinear time series analysis of the human electrocardiogram, *European Journal of Physics*, 26, 757-768.
[20] Schreiber T. [1993], Extremely simple nonlinear noise-reduction method, *Physical Review E*, 47.4, 2401-2404.
[21] Schreiber T. [1999], Interdisciplinary application of nonlinear time series methods, *Physics Reports*, 308, 1-64.
[22] Shintani M., Linton O. [2001], Is There Chaos in the World Economy? A Nonparametric Test Using Consistent Standard Errors, working paper, Vanderbilt University Nashville.
[23] Sivakumar B., Phoon K.-K., Liong S.-Y., Liaw C.-Y. [1999], A systematic approach to noise reduction in chaotic hydrological time series, *Journal of Hydrology*, 219, 103-135.
[24] Takens F. [1981], Detecting Strange Attractors in Turbulence, in: *Dynamical Systems and Turbulence*, Rand D., Young L., eds., Springer-Verlag, 366-381.
[25] Zeng X., Pielke R.A., Eykholt R. [1992], Extracting Lyapunov exponents from short time series of low precision, *Modern Physics Letters B*, 6, 55-75.

WPŁYW REDUKCJI SZUMU LOSOWEGO METODĄ SCHREIBERA NA IDENTYFIKACJĘ SYSTEMU GENERUJĄCEGO DANE. ANALIZA SYMULACYJNA

S t r e s z c z e n i e

Jednym ze sposobów ograniczenia negatywnego wpływu obecności szumu losowego na analizę rzeczywistych szeregów czasowych jest stosowanie metod redukcji szumu. Prezentowane w literaturze przedmiotu rezultaty zastosowania takich procedur w procesie identyfikacji nieliniowości i chaosu są zachęcające. Jedną z metod redukcji szumu jest metoda Schreibera, która, jak wykazano, prowadzi do efektywnej redukcji szumu losowego dodanego do danych wygenerowanych z systemów deterministycznych o dynamice chaotycznej. Jednakże w przypadku danych rzeczywistych, badacz zwykle pozbawiony jest wiedzy, czy system generujący rzeczywiście jest deterministyczny. Istnieje więc ryzyko, że redukcji szumu zostaną wówczas poddane dane losowe. W niniejszym artykule wykazano, iż w sytuacji, gdy brak jest wyraźnych podstaw do stwierdzenia, że badany szereg pochodzi z systemu deterministycznego, metodę Schreibera należy stosować z dużą ostrożnością. Z przeprowadzonych symulacji, w których wykorzystano test BDS, miarę informacji wzajemnej oraz współczynnik korelacji liniowej Pearsona wynika bowiem,

że redukcja szumu może wprowadzić do analizowanych danych, zależności o charakterze nieliniowym. W efekcie szeregi losowe mogą zostać błędnie zidentyfikowane jako nieliniowe.

Słowa kluczowe: Identyfikacja nieliniowości, redukcja szumu losowego, metoda Schreibera, wskaźnik NRL, test BDS, miara informacji wzajemnej, bootstrap

IMPACT OF THE SCHREIBER NOISE REDUCTION METHOD ON DGP IDENTIFICATION. SIMULATION ANALYSIS

S u m m a r y

A presence of a noise is typical for real-world data. In order to avoid its negative impact on methods of time series analysis, noise reduction procedures may be used. The achieved results of an application of such procedures in identification of chaos or nonlinearity seem to be encouraging. One of the noise reduction methods is the Schreiber method, which, as it has been shown, is able to effectively reduce a noise added to time series generated by deterministic systems with chaotic dynamics. However, while analyzing real-world data, a researcher usually cannot be sure if the generating system is deterministic. Therefore, there is a risk that a noise reduction method will be applied to random data. In this paper, it has been shown that in situations where there in no clear evidence that investigated data are generated by a deterministic system, the Schreiber noise reduction method may negatively affect identification of time series. In the simulation carried out in this paper, the BDS test, the mutual information measure and the Pearson autocorrelation coefficient were used. The research has shown that an application of the Schreiber method may introduce spurious nonlinear dependencies to investigated data. As a result, random series may be misidentified as nonlinear.

Keywords: Nonlinear time series, noise reduction, Schreiber method, NRL quantity, BDS test, mutual information, bootstrap

PRZEGLĄD STATYSTYCZNY
R. LVIII – ZESZYT 1-2 – 2011

KRZYSZTOF JAJUGA, TADEUSZ KUFEL, MAREK WALESIAK

SPRAWOZDANIE Z KONFERENCJI NAUKOWEJ NT. „KLASYFIKACJA I ANALIZA DANYCH – TEORIA I ZASTOSOWANIA"

W dniach 15-17 września 2010 roku w Hotelu Filmar w Toruniu odbyła się XIX Konferencja Naukowa Sekcji Klasyfikacji i Analizy Danych PTS (XXIV Konferencja Taksonomiczna) nt. Klasyfikacja i analiza danych – teoria i zastosowania, organizowana przez Sekcję Klasyfikacji i Analizy Danych Polskiego Towarzystwa Statystycznego, Katedrę Ekonometrii i Statystyki Uniwersytetu Mikołaja Kopernika w Toruniu oraz Katedrę Metod Ilościowych Wyższej Szkoły Bankowej w Toruniu. Przewodniczącym Komitetu Organizacyjnego Konferencji był dr hab. Tadeusz Kufel, prof. nadzw. UMK, natomiast sekretarzami dr Marcin Błażejowski i mgr Paweł Kufel.

Zakres tematyczny konferencji obejmował zagadnienia:

a) teoria (taksonomia, analiza dyskryminacyjna, metody porządkowania liniowego, metody statystycznej analizy wielowymiarowej, metody analizy zmiennych ciągłych, metody analizy zmiennych dyskretnych, metody analizy danych symbolicznych, metody graficzne),

b) zastosowania (analiza danych finansowych, analiza danych marketingowych, analiza danych przestrzennych, inne zastosowania analizy danych – medycyna, psychologia, archeologia, itd., aplikacje komputerowe metod statystycznych).

Zasadniczym celem konferencji SKAD była prezentacja osiągnięć i wymiana doświadczeń z zakresu teoretycznych i aplikacyjnych zagadnień klasyfikacji i analizy danych. Konferencja stanowi coroczne forum służące podsumowaniu obecnego stanu wiedzy, przedstawieniu i promocji dokonań nowatorskich oraz wskazaniu kierunków dalszych prac i badań.

W konferencji wzięło udział 112 osób. Byli to pracownicy oraz doktoranci następujących uczelni: Akademii Górniczo-Hutniczej w Krakowie, Politechniki Białostockiej, Politechniki Łódzkiej, Politechniki Opolskiej, Politechniki Szczecińskiej, Politechniki Wrocławskiej, Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie, Szkoły Głównej Handlowej w Warszawie, Uniwersytetu Ekonomicznego w Katowicach, Uniwersytetu Ekonomicznego w Krakowie, Uniwersytetu Ekonomicznego w Poznaniu, Uniwersytetu Ekonomicznego we Wrocławiu, Uniwersytetu Gdańskiego, Uniwersytetu im. Adama Mickiewicza w Poznaniu, Uniwersytetu Mikołaja Kopernika w Toruniu, Uniwersytetu Łódzkiego, Uniwersytetu Przyrodniczego w Poznaniu, Uniwersytetu Szczecińskiego, Uniwersytetu Warmińsko-Mazurskiego w Olsztynie, Uniwersytetu Warszawskiego, Wyższej Szkoły Zarządzania i Nauk Społecznych im. ks. Emila Szramka w Ty-

chach, Wyższej Szkoły Bankowej w Toruniu, Wyższej Szkoły Informatyki i Ekonomii w Olsztynie, Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie, a także przedstawiciele NBP, GUS oraz Wydawnictwa C.H. Beck.

W trakcie 3 sesji plenarnych oraz 12 sesji równoległych wygłoszono 60 referatów poświęconych aspektom teoretycznym i aplikacyjnym zagadnień klasyfikacji i analizy danych. Odbyła się również sesja plakatowa, na której zaprezentowano 23 plakaty.

Obradom w poszczególnych sesjach konferencji przewodniczyli: prof. dr hab. Krzysztof Jajuga, prof. dr hab. Eugeniusz Gatnar, prof. dr hab. Joanicjusz Nazarko, prof. dr hab. Marek Walesiak, prof. dr hab. Iwona Roeske-Słomka, prof. dr hab. Mirosław Krzyśko, prof. dr hab. Dorota Witkowska, prof. UEK dr hab. Andrzej Sokołowski, prof. UMK dr hab. Mariola Piłatowska, prof. UE dr hab. Jan Paradysz, prof. UE dr hab. Elżbieta Gołata, prof. dr hab. Jadwiga Suchecka, prof. dr hab. Bogdan Suchecki, prof. UE dr hab. Andrzej Bąk, prof. dr hab. Andrzej St. Barczak.

Teksty referatów przygotowane w formie recenzowanych artykułów naukowych stanowią zawartość przygotowywanej do druku publikacji z serii Taksonomia nr 18 (w ramach Prac Naukowych Uniwersytetu Ekonomicznego we Wrocławiu).

Zaprezentowano następujące referaty:

1. Stanisława Bartosiewicz (Wyższa Szkoła Bankowa we Wrocławiu), Opowieść o skutkach subiektywizmu w analizie wielowymiarowej

   Artykuł traktuje o weryfikacji tez, że subiektywizm „merytoryczny" (subiektywny u każdego badacza wybór cech opisujących zjawiska złożone w celu opracowania rankingu badanych obiektów) oraz subiektywizm „metodyczny" (subiektywny u każdego badacza wybór metod manipulacji na zbiorach cech) powodują trudności u odbiorców badań dotyczące wyboru właściwego rankingu obiektów z pośród różnych efektów uzyskanych przez różnych autorów badań. Główny wniosek z weryfikacji to: kłopoty odbiorców badań wywołuje jedynie zróżnicowanie wybranego subiektywnie zbioru cech.
2. Eugeniusz Gatnar (Uniwersytet Ekonomiczny w Katowicach), Statystyka i prawda. Uwagi o kryzysie finansowym

   W artykule przedstawiono spojrzenie na obecny kryzys finansowy, jako na kryzys zasad statystyki i prawdy. Statystyka jest zbiorem reguł i metod postępowania w celu odkrycia wiedzy o zjawiskach ekonomicznych, lub szerzej – społecznych, oraz ich zachowaniu. Kreatywna księgowość, transakcje za pomocą m.in. instrumentów typu FX swap były wykorzystywane po to, by ukryć prawdę o wysokości deficytu budżetowego oraz długu publicznego. Co więcej banki, fundusze oraz agencje ratingowe zrobiły wszystko, by zmaksymalizować ich zyski, na przykład oferując kredyty osobom nie posiadającym zdolności kredytowej, nazywanym NINJA. Również metody statystyczne, takie jak na przykład model VaR, były wykorzystywane niezgodnie z zasadami, do oceny wielkości rezerw na wypadek bankructwa.
3. Magdalena Osińska, Michał Bernard Pietrzak, Mirosława Żurek (Uniwersytet Mikołaja Kopernika w Toruniu), Wykorzystanie modeli równań strukturalnych do opisu mechanizmów podejmowania decyzji na rynku kapitałowym

W ostatnich latach, szczególnie w okresie kryzysu, obserwowany jest silny wzrost zainteresowania dziedziną finansów behawioralnych. Wskazywane jest coraz większe znaczenie wpływu inklinacji behawioralnych w procesie podejmowania decyzji inwestycyjnych, przy czym czynniki behawioralne rozumiane są tutaj jako zmienne istotnie zakłócające podjęcie prawidłowej decyzji w procesie inwestowania. Artykuł dotyczy weryfikacji wybranych elementów teorii finansów behawioralnych, na podstawie badania ankietowego, z wykorzystaniem modelu równań strukturalnych (SEM). Celem artykułu jest identyfikacja i opis inklinacji behawioralnych oraz weryfikacja postawionej hipotezy badawczej o ich wpływie na skłonność do ryzyka inwestorów indywidualnych.

4. Dorota Witkowska (Szkoła Główna Gospodarstwa Wiejskiego w Warszawie)), Próba oceny wpływu kapitału zagranicznego na efektywność chińskiego sektora bankowego za pomocą mierników syntetycznych

Coraz więcej zagranicznych banków inwestuje na chińskim rynku, pojawia się zatem pytanie czy obecność kapitału zagranicznego w chińskim sektorze bankowym wpływa na podniesienie jego efektywności. W celu odpowiedzi na to pytanie podjęto próbę aplikacji syntetycznych mierników rozwoju, skonstruowanych dla 72 chińskich banków na podstawie wskaźników finansowych za lata 2006 i 2007, do oceny ich efektywności. Przeprowadzono badanie zdolności dyskryminacyjnych wyznaczonych miar oraz wrażliwości i stabilności utworzonych za ich pomocą klas.

5. Marek Walesiak (Uniwersytet Ekonomiczny we Wrocławiu), Zastosowanie odległości GDM w analizie skupień na podstawie danych porządkowych

W artykule przedstawione dwa rozwiązania metodyczne (klasyczna analiza skupień i klasyfikacja spektralna) pozwalające na przeprowadzanie analizy skupień dla danych porządkowych z wykorzystaniem odległości GDM2. W części empirycznej zaprezentowane rozwiązania zilustrowano dla danych porządkowych z rynku nieruchomości z wykorzystaniem oprogramowania środowiska R.

6. Paweł Lula (Uniwersytet Ekonomiczny w Krakowie), Automatyczna analiza opinii konsumenckich

Celem artykułu jest próba klasyfikacji i scharakteryzowania metod automatyzacji analizy opinii konsumenckich oraz przedstawienie wyników prób ich zastosowania do analizy opinii polskojęzycznych. W pierwszej części pracy zaprezentowano klasyfikację opinii oraz metod ich analizy. Druga poświęcona jest systemom rozpoznającym ogólny charakter opinii. Część trzecia charakteryzuje systemy pozwalające na ocenę poszczególnych atrybutów ocenianych produktów. W referacie zaprezentowano proces budowy oraz ocenę dwóch różnych systemów analizy opinii konsumenckich. W ostatnim punkcie zamieszczono wnioski wynikające z przedstawionych badań.

7. Aleksandra Łuczak, Feliks Wysocki (Uniwersytet Przyrodniczy w Poznaniu), Zastosowanie analitycznego procesu sieciowego (ANP) w analizie SWOT jednostek administracyjnych

Przedstawiona metoda kwantyfikacji analizy SWOT z wykorzystaniem metody Saaty'ego analitycznego procesu sieciowego (ANP) jest kompleksową procedurą, która może być użyteczną w programowaniu rozwoju jednostek administracyjnych, szcze-

gólnie przy ocenie ich słabych i mocnych stron oraz szans i zagrożeń w ich otoczeniu. Ma ona przewagę nad metodami klasycznymi (opisowymi) ze względu na możliwość kwantyfikowania ważności czynników SWOT, a więc elementów o charakterze zarówno jakościowym, jak i ilościowym. Metoda ANP uwzględnia oprócz hierarchicznych powiązań pomiędzy elementami decyzyjnymi również interakcje między nimi – sprzężenia zwrotne. Może też być pomocna przy wyborze typu strategii rozwoju dla danej jednostki terytorialnej. Zagadnienie zilustrowano przykładem analizy SWOT gminy wiejskiej Babiak.

8. Beata Bal-Domańska (Uniwersytet Ekonomiczny we Wrocławiu), Konwergencja w regionach o różnym poziomie innowacyjności

Artykuł wpisuje się w nurt badań nad konwergencją. W pracy połączono dwa podejścia wykorzystywane w badaniach konwergencji: analizę sigma i beta, co pozwoliło na kompleksową oceną badanych procesów w regionach państw Unii Europejskiej (NUTS-2) w latach 1999-2007, które jako „całość" są jeszcze słabo rozpoznane. Procesy beta konwergencji warunkowej zidentyfikowano dzięki równaniu opartemu na strukturze modelu Mankiwa-Romera-Weila. Do oszacowania jego parametrów wykorzystano systemowy estymator Uogólnionej Metody Momentów [Arellano i Bover, Blundel i Bond]. Celem artykułu było rozpoznanie charakteru procesów konwergencji/dywergencji w regionach o różnym poziomie innowacyjności sektorowej.

9. Andrzej Dudek (Uniwersytet Ekonomiczny we Wrocławiu), Analiza danych symbolicznych w środowisku R. Podstawy metodologiczne i przykłady zastosowań

Analiza danych symbolicznych (Bock, Diday [2000], Billard, Diday [2006], Diday, Noirhome-Frature [2008]) to gałąź wielowymiarowej analizy statystycznej zajmująca się dużymi zbiorami danych (głównie pochodzącymi z komputerowych baz danych) zagregowanych w obiekty symboliczne, mogące zawierać dane w postaci liczb, tekstu, przedziałów liczbowych, zbiorów kategorii, zbiorów kategorii z wagami.

Artykuł zawiera opis autorskiego pakietu *SymbolicDA* służącego do analizy danych symbolicznych w popularnym środowisku **R** (http://www.r-project.org/) i składa się z dwóch części. Pierwsza jest próbą umiejscowienia podejścia symbolicznego w badaniach statystycznych ze szczególnym uwzględnieniem zastosowań ekonomicznych, druga zaś prezentuje funkcje pakietu wraz z przykładami ich zastosowań.

10. Katarzyna Kopczewska (Uniwersytet Warszawski), Modelowanie interakcji przestrzennych z wykorzystaniem programu R

Modele interakcji przestrzennych pozwalają ocenić znaczenie odległości w przepływie dóbr, osób, procesów, wiedzy, innowacji etc. Wykorzystywane powszechnie funkcje, w których przepływ ten tłumaczony jest odległością, mogą mieć postać wykładniczą, potęgową lub wielomianową, zaś ich estymacja możliwa jest przy wykorzystaniu metod klasycznych lub przestrzennych, w których dobór macierzy wag przestrzennych może być według różnych kryteriów sąsiedztwa. Celem artykułu jest analiza porównawcza powyższych modeli na danych dla gmin i powiatów, z uwzględnieniem oceny jakości dopasowania przez SRMSE czy zysk informacyjny, a także prezentacja metod aplikacji w programie R.

11. Christian Lis (Uniwersytet Szczeciński), Analiza porównawcza poziomu społeczno-gospodarczego rozwoju krajów Unii Europejskiej z uwzględnieniem paradygmatu gender equality

Celem artykułu jest ocena zróżnicowania poziomu rozwoju społeczno-gospodarczego krajów Unii Europejskiej z uwzględnieniem paradygmatu *gender equality*. W artykule zaprezentowane zostały w ocenie zróżnicowania poziomu rozwoju takie narzędzia wielowymiarowej analizy porównawczej jak taksonomiczny miernik poziomu rozwoju (TMPR), uogólniona miara rozwoju (GDM) i analiza skupień. Autor po raz pierwszy w swoich badaniach porównawczych rozszerzył zakres analizy o istotne czynniki leżące w sferze zainteresowań tzw. *gender mainstreaming*. W artykule wykorzystano m.in. komponenty GEM (*Gender Empowerment Measure*) i GDI (*Gender-related Development Index*).

12. Alicja Grześkowiak, Agnieszka Stanimir (Uniwersytet Ekonomiczny we Wrocławiu), Wielowymiarowa analiza wybranych aspektów postaw życiowych młodszych i starszych pokoleń Europejczyków

W pracy zaprezentowano porównanie postaw życiowych osób młodych (do 25 roku życia) oraz starszych (60+) zamieszkujących kraje europejskie. Ze względu na charakter dostępnych danych zastosowano klasyfikację dynamiczną opartą na miarach pozycyjnych oraz analizę czynnikową. Przeprowadzona analiza pozwoliła wskazać różnice w stopniu zaangażowania społecznego ludzi z różnych grup wiekowych oraz wyodrębnić główne elementy systemów wartości deklarowanych przez starsze i młodsze pokolenia.

13. Barbara Batóg, Magdalena Mojsiewicz (Uniwersytet Szczeciński), Katarzyna Wawrzyniak (Zachodniopomorski Uniwersytet Technologiczny w Szczecinie), Segmentacja gospodarstw domowych ze względu na popyt potencjalny i zrealizowany na rynku ubezpieczeń życiowych w Polsce

W artykule dokonano segmentacji gospodarstw domowych ze względu na popyt potencjalny i zrealizowany na rynku ubezpieczeń życiowych w Polsce z wykorzystaniem drzew regresyjnych i klasyfikacyjnych. Popyt potencjalny zdefiniowano jako zadeklarowaną miesięczną składkę na ubezpieczenia życiowe, natomiast popyt zrealizowany jako łączną liczbę polis na życie posiadanych przez gospodarstwo domowe. Dla każdego rodzaju popytu wydzielono segmenty gospodarstw domowych na podstawie cech demograficzno-ekonomicznych oraz preferencji dotyczących ubezpieczeń na życie. Chociaż zbiór zmiennych objaśniających był taki sam dla obu rodzajów popytu, to ich ranking ważności był inny w drzewach klasyfikacyjnych niż w drzewach regresyjnych.

14. Aleksandra Szlachcińska (Oddział Kliniczny Chirurgii Klatki Piersiowej i Rehabilitacji Oddechowej WSS im. M. Kopernika w Łodzi), Anna Witaszczyk, Małgorzata Misztal (Uniwersytet Łódzki), O zastosowaniu metody wiązania modeli do poprawy dokładności klasyfikacji pacjentów z pojedynczym cieniem okrągłym płuca

Metoda wiązania modeli (*bundling*) została zaproponowana przez Hothorna [2003] jako modyfikacja metody *bagging* [Breiman 1996]. Polega ona na wykorzystaniu do-

datkowych modeli, innych klas niż drzewa klasyfikacyjne, budowanych na podstawie zbioru OOB (*out-of-bag*), zawierającego obserwacje spoza aktualnej próby bootstrapowej. Na podstawie tych modeli dokonuje się predykcji dla obserwacji w próbie bootstrapowej a następnie wyniki predykcji traktuje się jako dodatkowe zmienne objaśniające przy budowie drzewa klasyfikacyjnego. W artykule przedstawiono wyniki wykorzystania metody wiązania modeli do poprawy dokładności klasyfikacji pacjentów z pojedynczym cieniem okrągłym płuca.

15. Iwona Bąk, Agnieszka Sompolska-Rzechuła (Zachodniopomorski Uniwersytet Technologiczny w Szczecinie), Wykorzystanie analizy log-liniowej do wyboru czynników determinujących wyjazdy turystyczne wybranej kategorii gospodarstw domowych

Celem artykułu jest próba wyodrębnienia zmiennych, które wpływają na podjęcie decyzji o wyjeździe turystycznym w gospodarstwach domowych emerytów i rencistów. Informacje dotyczące aktywności turystycznej gospodarstw domowych emerytów i rencistów zaczerpnięto z badań ankietowych „Turystyka i wypoczynek w gospodarstwach domowych" przeprowadzonych przez GUS w 2005 roku. Ponieważ w badaniu wzięto pod uwagę zmienne kategoryzacyjne, zatem do wyboru optymalnego zbioru czynników decydujących o wyjeździe turystycznym wykorzystano analizę log-liniową.

16. Elżbieta Gołata, Grażyna Dehnel, Hanna Gruchociak (Uniwersytet Ekonomiczny w Poznaniu), Próba wyodrębnienia stref wpływu wielkich miast w świetle badania dojazdów do pracy

W opracowaniu podjęto próbę delimitacji stref wpływu wielkich miast bazując na analogii do modelu Thünena, konstrukcji modeli interakcji przestrzennych ze szczególnym uwzględnieniem przemian ekonomicznych, demograficznych i społecznych obserwowanych w okresie transformacji w przestrzeni obszarów metropolitalnych. Podstawą analizy będą statystyki globalne Moran i Gary'ego oraz statystyka lokalna Morana. Celem opracowania będzie ponadto charakterystyka dojazdów do pracy oraz wskazanie możliwości ich wykorzystania w analizie rynku pracy w przekroju regionalnym i lokalnym. W artykule wykorzystano wyniki unikatowego badania przeprowadzonego w Urzędzie Statystycznym w Poznaniu, które dotyczyło przepływów związanych z zatrudnieniem.

17. Jan Paradysz, Karolina Paradysz (Uniwersytet Ekonomiczny w Poznaniu), Benchmarking w statystyce małych obszarów

Badania reprezentacyjne dostarczają wiarygodnych szacunków nie tylko dla całej zbiorowości lecz dla wielu subpopulacji, takich jak np. regiony, powiaty. Wielkość prób w małych domenach, szczególnie w małych przestrzennie obszarach rzadko kiedy pozwala na szacunki bezpośrednie. Zatem estymacja pośrednia odgrywa wiodącą rolę w zaspokajaniu stale rosnących potrzeb na wiarygodną statystykę nawet, gdy dostępne są tylko bardzo małe próby. Dokonując estymacji dla małych obszarów niezbędne jest „pożyczanie mocy" celem zwiększenia dla nich precyzji szacunku. W zależności od modelu estymacji oraz zmiennych wspomagających dla każdego małego obszaru uzyskujemy wiele szacunków. W związku z tym powstają problemy z wyborem właści-

wego szacunku. Benchmarking pozwala na właściwy wybór odpowiedniego szacunku. Zaproponowano trzy rodzaje kryteriów: poziomu, porządku i dystansu.

18. Mirosława Sztemberg-Lewandowska (Uniwersytet Ekonomiczny we Wrocławiu), Poziom edukacji w Europie z wykorzystaniem modelu krzywych rozwojowych

SEM inaczej nazywane LISREL są ogólną, głównie liniową, wielowymiarową techniką statystyczną. Jest ona bardziej konfirmacyjna niż eksploracyjna, czyli wykorzystuje się ją do sprawdzania dopasowania określonego modelu do danych, a nie do budowy pasującego modelu. Jednym ze złożonych modeli strukturalnych jest model krzywych rozwojowych, który służy do analizy zmiany wartości poszczególnych zmiennych w czasie. Model ten zakłada, że zmiana jest procesem ciągłym scharakteryzowanym przez zmienne, których realizacje różnią się między obiektami. Celem artykułu jest analiza procesu zmian w poziomie edukacji państw europejskich z wykorzystaniem krzywych rozwojowych. Interesującym jest też pokazanie Polski na tle Europy. Walor pracy polega na wykorzystaniu modeli równań strukturalnych do budowy latentnych krzywych rozwojowych opisanych nie tylko funkcją liniową, wielomianową, ale także wykładniczą.

19. Ewa Witek (Uniwersytet Ekonomiczny w Katowicach), Wykorzystanie mieszanek rozkładów Poissona do oceny liczby przyznanych patentów w krajach UE

W artykule przedstawiono zastosowanie mieszanek warunkowych rozkładów Poissona w regresji. Mieszanki tych rozkładów stosowane są wówczas gdy zbiór obserwacji charakteryzuje się nadmiernym rozproszeniem, będącym wynikiem np. pominięcia jednej z ważnych zmiennych objaśniających. Celem referatu jest zbadanie wpływu wydatków na badania i rozwój, na liczbę przydzielonych patentów w krajach Unii Europejskiej.

20. Andrzej Bąk, Tomasz Bartłomowicz (Uniwersytet Ekonomiczny we Wrocławiu), Implementacja klasycznej metody conjoint analysis w pakiecie conjoint programu R

Celem artykułu jest prezentacja pakietu conjoint opracowanego dla programu R, który należy obecnie do najważniejszych programów niekomercyjnych (oferowanych na zasadach licencji GNU) w zakresie analizy statystycznej i ekonometrycznej. Przedstawione zostały funkcje pakietu conjoint oraz zastosowania w badaniach marketingowych. Pakiet zawiera implementację klasycznej metody *conjoint analysis*, ale przygotowywane są rozszerzenia umożliwiające uwzględnienie modeli wyborów dyskretnych. Sposób użycia wybranych funkcji pakietu zilustrowano przykładami z zakresu badań preferencji.

21. Dorota Rozmus (Uniwersytet Ekonomiczny w Katowicach), Porównanie stabilności zagregowanych algorytmów taksonomicznych opartych na macierzy współwystąpień

Podejście zagregowane (wielomodelowe) dotychczas z dużym powodzeniem stosowane było w dyskryminacji w celu podniesienia dokładności klasyfikacji. W ostatnich latach analogiczne propozycje pojawiły się w taksonomii, aby zapewnić większą poprawność i stabilność wyników grupowania. Stabilność algorytmu taksonomicznego

w odniesieniu do niewielkich zmian w zbiorze danych, czy też parametrów algorytmu jest pożądaną cechą algorytmu. Głównym celem tego referatu jest porównanie stabilności zagregowanych algorytmów taksonomicznych opartych na macierzy współwystąpień oraz zbadanie relacji, jakie zachodzą między stabilnością a dokładnością.

22. Aneta Rybicka (Uniwersytet Ekonomiczny we Wrocławiu), Modele klas ukrytych w metodach wyborów dyskretnych

W badaniu preferencji wyrażonych wykorzystujemy m. in. metody wyborów dyskretnych reprezentujących podejście dekompozycyjne. W związku z tym, iż użyteczności cząstkowe oraz całkowite oszacowywane są na poziome zagregowanym, niemożliwe jest bezpośrednie przeprowadzenie segmentacji konsumentów. W celu oszacowania, w metodach wyborów dyskretnych, użyteczności na poziomie segmentowym, wykorzystujemy modele klas ukrytych. W artykule przedstawiono rodzaje modeli klas ukrytych, metodę estymacji parametrów, oprogramowanie komputerowe oraz przykłady zastosowań.

23. Tomasz Klimanek, Marcin Szymkowiak (Uniwersytet Ekonomiczny w Poznaniu), Taksonomiczne aspekty estymacji pośredniej uwzględniającej korelację przestrzenną

Głównym celem artykułu jest prezentacja metod i technik estymacji pośredniej uwzględniających przestrzeń. Wykorzystując dane z Narodowego Spisu Rolnego 2002 oraz mierniki autokorelacji przestrzennej, autorzy podejmują próbę oceny obciążenia estymatora EBLUP oraz estymatora EBLUP uwzględniającego korelację przestrzenną SEBLUP. Wyniki przeprowadzonych symulacji wskazują, że wykorzystanie informacji *a priori* o występowaniu bądź braku autokorelacji przestrzennej badanego zjawiska może w znaczący sposób poprawić jakość uzyskanych oszacowań (obciążenie estymatora).

24. Anna Czapkiewicz, Beata Basiura (AGH w Krakowie), Grupowanie indeksów światowych z uwzględnieniem przesunięć czasowych na podstawie modeli Copula-GARCH

W pracy zaprezentowana została próba pogrupowania danych, którymi są dzienne stopy zwrotu 42 indeksów światowych. Jako miarę powiązań między poszczególnymi indeksami przyjęto współczynnik korelacji, który jest parametrem funkcji połączeń t-Studenta. W oparciu o ten współczynnik zdefiniowano miarę odległości, pozwalającą utworzyć podział na grupy taksonomiczne. Celem badania jest określenie czy istnieje wpływ przesunięcia czasowego na wyniki grupowania.

25. Mariola Chrzanowska (Szkoła Główna Gospodarstwa Wiejskiego w Warszawie), Propozycja doboru cech w wielowymiarowej analizie porównawczej na przykładzie giełd Europy Środkowo-Wschodniej

Postępująca globalizacja umożliwia dostęp do coraz większej liczby informacji. Wpływa również na częstotliwość pozyskiwania danych. Niektóre zjawiska notowane są w cyklu miesięcznym, dziennym a nawet minutowym. Podczas badań takich procesów pojawia się problem wyboru odpowiednich, właściwych informacji. W pracy

zaprezentowano propozycję rozwiązania tego problemu na przykładzie giełd Europy Środkowo-Wschodniej.

26. Małgorzata Misztal (Uniwersytet Łódzki), Próba oceny wpływu wybranych metod imputacji danych na wyniki klasyfikacji obiektów z wykorzystaniem drzew klasyfikacyjnych

W praktycznych zastosowaniach metod statystycznych często pojawia się problem występowania w zbiorach danych brakujących wartości. W takiej sytuacji wymienić można trzy sposoby postępowania: (1) odrzucenie obiektów z wartościami brakującymi, (2) wykorzystanie algorytmu uczącego do rozwiązania problemu brakujących wartości w fazie uczenia, (3) imputację brakujących wartości przed zastosowaniem algorytmu uczącego. Celem głównym pracy jest ocena wpływu metod postępowania w sytuacji występowania braków danych na wyniki klasyfikacji obiektów za pomocą drzew klasyfikacyjnych.

27. Barbara Batóg, Jacek Batóg (Uniwersytet Szczeciński), Klasyfikacja największych polskich przedsiębiorstw według wydajności pracy w ujęciu dynamicznym

W artykule dokonano analizy zmian wydajności pracy największych polskich przedsiębiorstw z wykorzystaniem danych panelowych w ujęciu sektorowym w latach 2004-2008. W oparciu o utworzone tabele kontyngencji oraz wybrane miary heterogeniczności dokonano podziału wszystkich badanych obiektów na podzbiory jednorodne. Tak uporządkowane dane panelowe posłużyły do analizy i oceny zachodzących zmian w czasie i w przestrzeni zarówno w ogólnym poziomie wydajności pracy, jak i zmian przynależności poszczególnych obiektów do określonych klas wydajności (macierze przejścia).

28. Aleksandra Witkowska, Marek Witkowski (Uniwersytet Ekonomiczny w Poznaniu), Zmienna syntetyczna z medianą w ocenie kondycji finansowej banków spółdzielczych

W pracy podjęto próbę zastosowania do oceny stanu kondycji finansowej podmiotów gospodarczych zmiennej syntetycznej. Jest to tzw. zmienna syntetyczna z medianą. Przedmiotem zainteresowania, jako obiekty badania, były banki spółdzielcze, które charakteryzują się bardzo zróżnicowaną wielkością, a co za tym idzie i skalą działalności. Sądzimy w związku z tym, że zastosowanie zmiennej syntetycznej z medianą ma w tej sytuacji swoje uzasadnienie. Pokrycie informacyjne dla prowadzonych rozważań uzyskaliśmy wykorzystując dane pochodzące z banków spółdzielczych należących do jednego ze zrzeszeń tych banków. Dane te pochodziły ze sprawozdań finansowych tych banków i dotyczyły lat 2004-2007

29. Artur Zaborski (Uniwersytet Ekonomiczny we Wrocławiu), Zastosowanie algorytmu SMACOF do badan opartych na prostokątnej macierzy preferencji

SMACOF jest strategią skalowania wielowymiarowego wykorzystującą metodę majoryzacji, która aproksymuje w kolejnych cyklach iteracyjnych minimalne wartości funkcji STRESS. Celem artykułu jest prezentacja metodologii skalowania wielowymiarowego za pomocą dostępnego w środowisku R algorytmu SMACOF i jego mody-

fikacji na potrzeby prostokątnej macierzy preferencji. Na zakończenie zaprezentowano przykład, w którym wykorzystano funkcję smasofRect pakietu smacof.

30. Kamila Migdał Najman (Uniwersytet Gdański), Analiza porównawcza samouczących się sieci neuronowych typu SOM i GNG w poszukiwaniu reguł asocjacyjnych

W artykule autorka dokonuje próby porównania dwóch metod analizy skupień opartych na nienadzorowanym uczeniu sieci neuronowych SOM i GNG w poszukiwaniu reguł asocjacyjnych. Autorka weryfikuje potencjał samouczących się sieci w poszukiwaniu wzorców zakupowych klientów na dwóch zbiorach danych umownych. Wykryte reguły asocjacyjne uzyskane w oparciu o sieć SOM i GNG prezentuje wizualnie ma wykresach sieciowych.

31. Krzysztof Najman (Uniwersytet Gdański), Propozycja algorytmu samouczenia się sieci neuronowych typu GNG ze zmiennym krokiem uczenia

Jednym z kluczowych parametrów procesu samouczenia się sieci neuronowych typu GNG jest szybkość zmiany pozycji w przestrzeni neuronu uczącego się i najbliższego połączonego z nim neuronu. Zależy ona od lokalnego błędu kwantyzacji i stałej nazywanej krokiem uczenia. Stała wartość kroku uczenia w szczególności niepotrzebnie zwalnia proces samouczenia się w początkowej jego fazie. W artykule proponuje się modyfikację algorytmu wprowadzając zmienny krok uczenia oparty na liniowej funkcji iteracji między kolejnymi fazami wstawiania nowego neuronu do sieci. Przeprowadzone rozważania teoretyczne i eksperymenty symulacyjne potwierdzają zasadność proponowanej zmiany.

32. Justyna Wilk (Uniwersytet Ekonomiczny we Wrocławiu), Klasyfikacja urzędów gminnych w zakresie e-administracji

E-administracja polega na udostępnieniu interesantom (mieszkańcom i przedsiębiorstwom) usług publicznych za pośrednictwem Internetu. Istotną rolę w rozwoju e-administracji odgrywa prowadzenie badań porównawczych w zakresie e-usług oferowanych przez instytucje administracji publicznej. Celem artykułu jest klasyfikacja urzędów gminnych województwa dolnośląskiego w zakresie e-usług. Procedura badawcza obejmowała dwa etapy: agregację danych i analizę skupień. W wyniku zastosowanej procedury wyodrębniono 6 klas gmin zróżnicowanych ze względu na zakres informacji i formy komunikacji oferowane na stronach internetowych.

33. Julita Stańczuk, Patrycja Trojczak-Golonka (Zachodniopomorski Uniwersytet Technologiczny w Szczecinie), Analiza wpływu pogorszenia jakości informacji opisowej na wynik klasyfikacji obiektów ekonomicznych z wykorzystaniem sieci neuronowych

Celem artykułu jest przedstawienie znaczenia informacji opisowej dla klasyfikacji przedsiębiorstw notowanych na GPW w Warszawie, a dokładniej możliwości wystąpienia braków, danych zaszumionych, czy celowej redukcji liczby zmiennych. Istotne jest to, w jaki sposób pogorszenie tej jakości wpływa na efektywność klasyfikacji, a więc przede wszystkim na liczbę przedsiębiorstw poprawnie zaklasyfikowanych do poszczególnych grup (z wykorzystaniem ratingu). Próbę badawczą tworzą przedsiębiorstwa notowane na GPW, dane natomiast pochodzą z ich sprawozdań finansowych.

W badaniu wykorzystano sieci neuronowe umożliwiające m.in. klasyfikację obiektów. Posłużono się wcześniejszymi badaniami do porównania otrzymanych wyników.

34. Robert Kapłon (Politechnika Wrocławska), O pewnych sposobach uwzględnienia niejednorodności obserwacji w modelach częstości zakupów

Dane dotyczące częstości zakupów są obecnie łatwo dostępne za sprawą systemów transakcyjnych, które je zbierają i przechowują. Dlatego potrzebne są odpowiednie narzędzia pozwalające na ich analizę. Często wykorzystywany model Poissona, ze względu na niejednorodność danych, jest nieodpowiedni. W tej sytuacji należy poszukiwać bardziej złożonych i zarazem bardziej wiarygodnych modeli. W pracy zaprezentowano dwa konkurencyjne modele pozwalające na uchwycenie niejednorodności: mieszanki rozkładów Poissona oraz mieszane modele Poissona. Wychodząc natomiast od przesłanek teoretycznych i empirycznych, wskazano na podobieństwa i różnice między nimi.

35. Bartłomiej Jefmański (Uniwersytet Ekonomiczny we Wrocławiu), Zastosowanie rozmytej funkcji regresji w ocenie poziomu satysfakcji z usług

Model regresji rozmytej zaproponowany przez Tanakę a określany mianem posybilistycznej funkcji regresji jest nieparametryczną metodą użyteczną w estymacji rozmytych zależności między zmiennymi. Jego celem jest minimalizacja stopnia rozmycia związku między zmiennymi poprzez rozwiązanie zadania programowania matematycznego. Celem artykułu jest zaprezentowanie możliwości wykorzystania regresji rozmytej w badaniach satysfakcji konsumentów a w szczególności szacowania ich poziomu satysfakcji. Podstawy teoretyczne proponowanego podejścia zaprezentowano na przykładzie analizy wyników otrzymanych z badania satysfakcji studentów jednej z niepublicznych szkół wyższych.

36. Michał Trzęsiok (Uniwersytet Ekonomiczny w Katowicach), Gdzie jest trzeci świat? – analiza taksonomiczna

W artykule przedstawiono próbę wielowymiarowego spojrzenia na problem identyfikacji krajów trzeciego świata z wykorzystaniem modeli taksonomicznych i porównanie wyników klasyfikacji państw otrzymanych z analizy danych z ONZ za 2008 oraz 1973 rok. Celem artykułu jest weryfikacja hipotezy sformułowanej przez prof. Hansa Roslinga głoszącej, że używany niegdyś podział państw na kraje rozwinięte i kraje trzeciego świata staje się coraz mniej uprawniony wobec znaczących zmian w sytuacji społeczno-gospodarczej i szybkiego rozwoju licznej grupy państw Ameryki Południowej, Azji i Afryki.

37. Joanna Trzęsiok (Uniwersytet Ekonomiczny w Katowicach), Przegląd metod regularyzacji w zagadnieniach regresji nieparametrycznej

Wiele nieparametrycznych funkcji regresji, w trakcie wykonywania algorytmu, jest systematycznie poprawianych, tak by końcowy model charakteryzował się jak najlepszym dopasowaniem do danych ze zbioru uczącego. W efekcie otrzymujemy modele o niskich wartościach błędów resubstytucji i wysokiej złożoności, które jednak charakteryzują się niewielką zdolnością uogólniania rozumianą jako zdolność poprawnej predykcji na nowych obiektach. Zachodzi potrzeba przeciwdziałania temu zjawisku. Proces uproszczenia postaci modelu przy jednoczesnym kontrolowaniu jego dopaso-

wania nazywamy regularyzacją. W artykule przedstawione i porównane zostały techniki regularyzacji wykorzystywane w nieparametrycznych metodach regresji.

38. Katarzyna Wójcik (Uniwersytet Ekonomiczny w Krakowie), Analiza porównawcza miar podobieństwa tekstów

Zasadniczym celem niniejszej pracy jest próba oceny przydatności znanych z literatury miar podobieństwa tekstów. W kolejnych punktach artykułu przedstawione zostały najpierw dokumenty tekstowe, które były porównywane w badaniu, a następnie wybrane miary podobieństwa oparte ma macierzy częstości wykorzystane do porównania dokumentów. W dalszej części zaprezentowane zostały wyniki przeprowadzonej analizy symulacyjnej opisanych wcześniej miar. Na tej podstawie została podjęta próba oceny przydatności tych miar.

39. Małgorzata Gliwa (Uniwersytet Ekonomiczny w Katowicach), Wpływ metody dyskretyzacji na jakość klasyfikacji

Główny cel artykułu to porównanie wielkości błędów klasyfikacji dla modeli dyskryminacyjnych zbudowanych dla zbiorów danych przed dyskretyzacją i po dyskretyzacji. Jako metodę dyskryminacji zastosowano naiwny klasyfikator bayesowski. Modele budowano zarówno dla zbiorów danych przed dyskretyzacja, jak i po dyskretyzacji. Dyskretyzacji dokonano z wykorzystaniem metod bezkontekstowych (dyskretyzacja na równe przedziały i przedziały o równych liczebnościach) i kontekstowych (metoda ChiMerge i minimalizacji entropii). Obliczenia wykonano na podstawie autorskich procedur i funkcji zawartych w pakietach dprep, e1071, grDevices, infotheo oraz car programu **R**.

40. Mariusz Grabowski (Uniwersytet Ekonomiczny w Krakowie), Metoda przypisania dziedzinowej ontologii słów kluczowych do istniejących artykułów naukowych

Duża dostępność dokumentów tekstowych powoduje konieczność tworzenia mechanizmów ułatwiających ich wyszukiwanie i kategoryzację. W przypadku artykułów naukowych elementem ułatwiającym wymienione zadania jest lista swobodnie dobieranych słów kluczowych. Swobodny dobór słów kluczowych tylko częściowo ułatwia realizację wymienionych zadań, gdyż jego rezultatem jest pojawianie się synonimów i homonimów oraz nie pozwala on na tworzenie dziedzinowej hierarchii pojęciowej. Rozwiązaniem tego problemu może być zastosowanie dziedzinowej ontologii słów kluczowych. W artykule została zaprezentowana metoda bazująca na koncepcji podobieństwa semantycznego, pozwalająca na automatyczne przypisanie słów kluczowych ontologii ISRL (Information Systems Research Library) do istniejących artykułów, w których system ISRL nie był stosowany. Zaletą omawianej metody jest możliwość zdefiniowania ontologii słów kluczowych dla dowolnej dziedziny lub dyscypliny naukowej i przypisanie należących do niej słów kluczowych istniejącym już artykułom.

41. Jarosław Krajewski (Uniwersytet Mikołaja Kopernika w Toruniu), Wpływ transformacji danych na wyniki estymacji dynamicznego modelu czynnikowego za pomocą metody głównych składowych

Artykuł traktuje o dynamicznych modelach czynnikowych (DFM). Prezentuje ich ideę, zastosowanie metody głównych składowych do ich estymacji oraz zastosowa-

nie kryteriów informacyjnych Bai'a i Ng do specyfikacji liczby czynników. Głównym celem artykułu jest ukazanie wpływu transformacji danych makroekonomicznych na ostateczną postać DFM. W badaniu wykorzystano 62 zmienne makroekonomiczne w postaci szeregów czasowych o częstotliwości miesięcznej z okresu od stycznia 2002 do marca 2009 roku. W wyniku szacowania DFM na różnych etapach transformacji danych uzyskano postacie DFM różniące się co do zawartych w nich zmiennych, jak i co do wartości parametrów stojących przy poszczególnych zmiennych.

42. Małgorzata Machowska-Szewczyk (Zachodniopomorski Uniwersytet Technologiczny w Szczecinie), Podobieństwa i różnice miedzy obiektami symbolicznymi a obiektami rozmytymi

Informacje zgromadzone w bazach danych często mają postać tekstu zapisanego w naturalnym języku. Informacje lingwistyczne są proste do przetwarzania dla człowieka, który posługuje się na co dzień takimi pojęciami, stanowi jednak dość poważną barierę dla komputera. Obiekty symboliczne oraz obiekty charakteryzowane przez cechy rozmyte umożliwiają przetworzenie danych wyrażonych lingwistycznie do postaci akceptowalnej przez komputer. Głównym celem artykułu jest analiza porównawcza, czyli przedstawienie podobieństw i różnic, jakie występują w opisie obiektów przedstawionych za pomocą zmiennych symbolicznych oraz zmiennych rozmytych.

43. Marcin Pełka (Uniwersytet Ekonomiczny we Wrocławiu), Podejście wielomodelowe z wykorzystaniem metody bagging w analizie danych symbolicznych na przykładzie metody k-najbliższych sąsiadów

W artykule przedstawiono podstawowe pojęcia związane z metodą bagging oraz metodą k-najbliższych sąsiadów (KNN) dla danych symbolicznych. Zaprezentowano w nim także możliwość zastosowania podejścia wielomodelowego bagging dla metody k-najbliższych sąsiadów dla danych symbolicznych. W części empirycznej przedstawiono zastosowanie podejścia wielomodelowego dla danych symbolicznych dla przykładowych zbiorów danych wygenerowanych za pomocą funkcji cluster.Gen z pakietu clusterSim w programie **R**.

44. Ewa Wędrowska (Uniwersytet Warmińsko-Mazurski w Olsztynie), Dywergencje Jensena-Shannona oraz Jensena-Rény'ego jako symetryczne wygładzenia miary Kullbacka-Leiblera

W artykule wskazano na możliwość wykorzystania do oceny stopnia rozbieżności struktur miar dywergencji. Zaprezentowano symetryczne „wygładzenie" dywergencji Kullbacka-Leiblera, jakim jest miara rozbieżności Jensena-Shannona, będąca funkcją entropii Shannona. Istotną zaletą dywergencji Jensena-Shannona jest możliwość uogólnienia jej do badania podobieństwa więcej niż dwóch struktur, a także możliwość uwzględnienia wag dla rozpatrywanych obiektów. Kolejną przedstawioną miarą dywergencji jest dywergencja Jensena-Rényi'ego stopnia α, która jest odpowiednio funkcją entropii Rényi'ego i jest uogólnieniem dywergencji Jensena-Shannona.

45. Justyna Brzezińska (Uniwersytet Ekonomiczny w Katowicach), Rozkład macierzy według wartości osobliwych (SVD) w analizie korespondencji

Wizualizacja wyników analizy korespondencji jest możliwa dzięki dekompozycji macierzy według wartości osobliwych (*Singular Value Decomposition*). Celem niniejszej pracy jest przedstawione różnych podejść i algorytmów metody *SVD* stosowanych w analizie korespondencji, dzięki którym możliwe jest zaprezentowanie kategorii badanych zmiennych w jednym układzie odniesienia. W literaturze wymieniane są algorytmy metody *SVD* według czterech podejść: Fishera [1940], Greenacre'a [1984], Andersena [1991] oraz Jobsona [1992]. Zaprezentowane zostaną także sposoby wyznaczania współrzędnych kategorii wierszowych oraz kolumnowych. Interesującym problemem jest porównanie wszystkich podejść, gdyż w literaturze wykorzystywane jest głównie podejście zaproponowane przez Greenacre'a. W pracy zaprezentowano graficzne przedstawienie wyników dla każdej z omawianej metody w programie **R**.

46. Tadeusz Kufel (Uniwersytet Mikołaja Kopernika w Toruniu), Marcin Błażejowski (Wyższa Szkoła Bankowa w Toruniu), Paweł Kufel (Uniwersytet Mikołaja Kopernika w Toruniu), Modelowanie zmiennych jakościowych i ograniczonych z wykorzystaniem oprogramowania GRETL

W artykule zostały przedstawione przykłady modelowania zmiennych ograniczonych, takich jak binarnych, dyskretnych, licznikowych, z wykorzystaniem estymacji logitowej dla zmiennej: dwumianowej, wielomianowej uporządkowanej, wielomianowej nieuporządkowanej, estymacji tobitowej i heckitowej (dla zmiennej uciętej), regresji Poissona i regresji ujemnej dwumianowej (dla zmiennej licznikowej). Całość zilustrowana została przykładami i procedurami estymacji z oprogramowania GRETL.

47. Grzegorz Kowalewski (Uniwersytet Ekonomiczny we Wrocławiu), Zastosowanie metod porządkowania liniowego w analizie koniunktury

Wśród metod służących do empirycznej analizy i prognozy koniunktury gospodarczej można wyróżnić metody wskaźnikowe, w których korzystamy z wielu zmiennych, czułych na zmiany koniunktury. Na podstawie tych zmiennych obliczane są wskaźniki zbiorcze (syntetyczne, złożone, ogólne), na podstawie których wnioskujemy o zmianach koniunktury gospodarczej. W artykule zaproponowano zastosowanie metod porządkowania liniowego w analizie koniunktury gospodarczej za pomocą metod wskaźnikowych.

48. Mariusz Kubus (Politechnika Opolska), Analiza metody LARS w problemie selekcji zmiennych w regresji

Selekcja zmiennych jest typowym zadaniem *data mining*, gdzie prowadzący analizę poszukuje interesujących i nieoczekiwanych relacji w danych bez wiedzy początkowej na temat badanego zjawiska. W liniowym modelu regresji, zamiast popularnej procedury krokowej czy też eliminacji zmiennych testem istotności współczynników, do selekcji zmiennych zastosować można metody iteracyjnej estymacji parametrów modelu (np. LARS Efrona i in. [2004]). Celem artykułu jest zbadanie zdolności metody LARS do identyfikowania zmiennych nieistotnych, szczególnie w przypadku, gdy zachodzą między nimi zależności liniowe. Dokonane też będzie porównanie z wybranymi metodami selekcji zmiennych.

49. Agnieszka Sompolska-Rzechuła, Iwona Bak (Zachodniopomorski Uniwersytet Technologiczny w Szczecinie), Zastosowanie klasyfikacji dynamicznej do oceny poziomu atrakcyjności turystycznej powiatów województwa zachodniopomorskiego

Celem artykułu jest próba określenia tendencji rozwoju powiatów pod względem poziomu atrakcyjności turystycznej oraz wydzielenie grup typologicznych badanych obiektów o podobnym poziomie dynamiki badanego zjawiska. Do klasyfikacji dynamicznej wykorzystano funkcje trendu oraz miary odległości między macierzami danych przekrojowo-czasowych. Analiza dotyczy lat 1999-2008.

50. Joanna Landmesser (Szkoła Główna Gospodarstwa Wiejskiego w Warszawie), Mikroekonometryczne metody pomiaru skuteczności udziału bezrobotnych w szkoleniach

Celem badania jest określenie wpływu szkoleń dla bezrobotnych na indywidualną długość czasu trwania w bezrobociu. Aby rozwiązać problem niewłaściwego doboru próby, wykorzystuje się dwie metody. W pierwszej stosujemy dobieranie, by znaleźć odpowiednią grupę kontrolną dla grupy szkolonych. W drugiej udział w szkoleniu jest instrumentalizowany za pomocą modelu probitowego. W dalszej kolejności są szacowane semiparametryczne modele hazardu. Badanie oparto o dane z Powiatowego Urzędu Pracy w Słupsku w Polsce.

51. Aleksandra Matuszewska-Janica (Szkoła Główna Gospodarstwa Wiejskiego w Warszawie), Nierówności w wynagrodzeniach kobiet i mężczyzn w krajach Unii Europejskiej: analiza conjoint

Z badań statystycznych wynika, że różnice w płacach kobiet i mężczyzn (*Gender Pay Gap*) w krajach UE przeciętnie wynoszą aż 18%. Jednakże są takie obszary na rynku pracy, gdzie różnice w płacach są znacznie mniejsze (lub znacznie większe) Celem pracy było wskazanie na scharakteryzowanie profili obszarów zawodowe (za pomocą rodzaju branży, wykształcenia i kategorii wiekowej), w których dysproporcje płacowe są najmniejsze i największe.

52. Iwona Foryś (Uniwersytet Szczeciński), Wielowymiarowa analiza cech mieszkań sprzedawanych na rynku warszawskim w badaniu czasu trwania oferty w systemie MLS

Przedmiotem analizy jest warszawski rynek mieszkaniowy. Negatywnie zweryfikowano hipotezę o stałości cech wpływających na płynność, niezależnie od sytuacji na rynku mieszkaniowym. W tym celu wykorzystano funkcję dyskryminacji. Wykazano, że funkcje dyskryminacyjne w sposób zadawalający klasyfikują transakcje do trzech zaproponowanych grup płynności: niska, wysoka i średnia.

53. Marcin Salamaga (Uniwersytet Ekonomiczny w Krakowie), Porównania międzynarodowe struktury bezpośrednich inwestycji zagranicznych z wykorzystaniem wielowymiarowej analizy statystycznej

Głównym celem artykułu jest porównanie struktury bezpośrednich inwestycji zagranicznych lokowanych przez kraje UE. Wykorzystując wielowymiarowe metody analizy danych przeprowadzono grupowania państw członkowskich UE ze względu na podobieństwo struktury branżowej i przestrzennej bezpośrednich inwestycji zagranicz-

nych. Zastosowanie analizy korespondencji oraz wybranych hierarchicznych metod grupowania pozwoliło na wyodrębnienie grup krajów o podobnych profilach inwestycyjnych i podobnym stopniu konkurencyjności inwestycyjnej.

54. Adam P. Balcerzak (Uniwersytet Mikołaja Kopernika w Toruniu), Taksonomiczna analiza jakości kapitału ludzkiego w Unii Europejskiej w latach 2002-2008

Artykuł prezentuje wielowymiarową analizę jakości kapitału ludzkiego w krajach UE. Zgodnie planem Europa 2020 kraje Wspólnoty muszą stale podnosić jakość kapitału ludzkiego, adekwatnie do wymogów konkurencyjnej gospodarki opartej na wiedzy. Kraje mogą działać w tym zakresie zgodnie z indywidualnymi strategiami rozwoju. Tym samym konieczne jest porównywanie ich wyników. Pozwoli to na wskazanie pozytywnych przykładów, które można wykorzystywać w procesie tworzenia strategii oraz przykładów negatywnych, których bezwzględnie należy unikać. Badanie dotyczy lat 2002-2008, zastosowano procedurę porządkowania liniowego bazującej na metodzie wzorca rozwoju Hellwiga, z zachowaniem stałego wzorca dla całego okresu oraz zmiennego wzorca dla danych Eurostatu.

55. Joanna Banaś, Małgorzata Machowska-Szewczyk (Zachodniopomorski Uniwersytet Technologiczny w Szczecinie), Zastosowanie uporządkowanego modelu probitowego do oceny połączenia Politechniki Szczecińskiej i Akademii Rolniczej w Szczecinie

W wyniku połączenia Politechniki Szczecińskiej oraz Akademii Rolniczej w Szczecinie w jedną uczelnię Zachodniopomorski Uniwersytet Technologiczny w Szczecinie (ZUT) w 2009 roku, nastąpiła zmiana organizacji pracy i nauczania w obu uczelniach. Podstawowym celem artykułu jest analiza opinii związanych z tym połączeniem w grupie pracowników, na podstawie wypełnianych dobrowolnie kwestionariuszy ankiety. Do oceny stopnia zadowolenia z połączenia obu uczelni zastosowano uporządkowany model probitowy, ponieważ prawie wszystkie zmienne przyjęte do opisu zjawiska mają charakter jakościowy, zaś zmienna objaśniana wyrażona jest w skali porządkowej.

56. Aneta Becker (Zachodniopomorski Uniwersytet Technologiczny w Szczecinie), Ranking województw pod względem stopnia wykorzystania technologii IC w przedsiębiorstwach

W artykule przedstawiono wyniki uporządkowania województw Polski pod względem wykorzystania technologii informacyjno-telekomunikacyjnych w przedsiębiorstwach, w 2009 r. W badaniach wykorzystano metodę UTA, która jest przykładem metody wielokryterialnego porządkowania wariantów decyzyjnych, oparta na dezagregacji informacji preferencyjnej za pomocą regresji porządkowej.

57. Katarzyna Dębkowska (Politechnika Białostocka), Wielowymiarowa analiza poziomu ubóstwa gmin województwa podlaskiego

Celem artykułu było zbadanie poziomu ubóstwa w powiatach województwa podlaskiego za pomocą metod statystyki wielowymiarowej, czyli metod porządkowania liniowego i nieliniowego. Uzyskane wyniki potwierdzają, że zaproponowane metody mogą być z powodzeniem wykorzystane do badania zjawiska ubóstwa. W rezultacie utworzono ranking powiatów od najmniej do najbardziej zagrożonych ubóstwem oraz

na pogrupowano powiaty na podobne pod względem zagrożenia ubóstwem. W pracy zaproponowano różne podejścia do klasyfikacji obiektów, w szczególności do szukania obiektów podobnych.

58. Anna Domagała (Uniwersytet Ekonomiczny w Poznaniu), Zastosowanie Data Envelopment Analysis do badania efektywności europejskich giełd papierów wartościowych: realizacja funkcji mobilizacji kapitału

Celem artykułu było zastosowanie Data Envelopment Analysis (DEA), metody z sukcesem wykorzystywanej do oceny efektywności w innych dziedzinach, w badaniu efektywności działania giełd papierów wartościowych. Metodę DEA wykorzystano do zbadania efektywności działania piętnastu europejskich giełd papierów wartościowych w zakresie realizacji funkcji mobilizacji kapitału w latach 2003-2005. Zastosowano nieradialny zorientowany na nakłady model SBM z nadefektywnością.

59. Iwona Staniec, Adam Depta (Politechnika Łódzka), Preferencje transportowe studentów łódzkich uczelni

Przedstawione w pracy rozważania dotyczą preferencji transportowych studentów łódzkich uczelni. Celem prowadzonych badań było pozyskanie wiedzy z jakiego środka komunikacji korzystają studenci dostając się na uczelnię oraz określenie odczuwanych przez nich wad i zalet wybranych środków komunikacji. W celu identyfikacji głównych czynników wpływających na preferencje transportowe studentów wykorzystano modele regresji logistycznej.

60. Patrycja Trojczak-Golonka, Julita Stańczuk (Zachodniopomorski Uniwersytet Technologiczny w Szczecinie), Specyfika danych finansowych i znaczenie etapu preprocessingu

Wykrywanie prawidłowości w badanych zjawiskach, interpretowanie ich za pomocą metod matematycznych i wyprowadzanie użytecznych wniosków to podstawa w analizie statystycznej. Sam zakres zastosowań badań statystycznych jest nieograniczony, a szczególnie użyteczny jest w opisie zjawisk masowych, takich jak np. przedsiębiorstwo. Choć pozornie wydawać się może, że dane finansowe są w swej istocie podobne do innych źródeł danych to w rzeczywistości nie jest to prawdą. W artykule zaprezentowano istotę danych finansowych i specyficzne cechy, które odróżniają je od innych typów danych. Omówione zostały również problemy związane z pozyskiwaniem i przygotowaniem ich do badań. Artykuł porządkuje wiedzę na temat danych finansowych i znaczenia ich swoistego charakteru dla wyników powszechnie prowadzonych analiz.

61. Julia Koralun-Bereźnicka (Uniwersytet Gdański), Efekt sektorowy i efekt wielkości podmiotu w kondycji finansowej przedsiębiorstw w Polsce

Celem badania jest ustalenie relatywnej ważności efektu sektorowego i efektu wielkości przedsiębiorstwa na jego kondycję finansową. Analiza obejmuje grupy przedsiębiorstw małych, średnich i dużych w trzynastu sektorach gospodarczych w Polsce w latach 2002-2007. Zmienne w postaci rocznych wskaźników finansowych obliczono na podstawie zagregowanych sprawozdań finansowych publikowanych w ramach bazy danych BACH. Metodologia badawcza obejmuje m.in. analizę skupień i procedurę

skalowania wielowymiarowego. Uzyskane wyniki świadczą o nieznacznej przewadze czynników sektorowych nad rozmiarem przedsiębiorstw w Polsce, co z kolei implikuje wyższość strategii dywersyfikacji opartych na przekrojach sektorowych, a nie na rozmiarach.

62. Izabela Kurzawa, Feliks Wysocki (Uniwersytet Przyrodniczy w Poznaniu), Zastosowanie modelu LA/AIDS do badania elastyczności cenowych popytu konsumpcyjnego

Celem pracy było zbadanie przydatności liniowej aproksymacji prawie idealnego systemu funkcji popytu LA/AIDS do wyznaczenia elastyczności cenowych popytu dla wybranych grup artykułów żywnościowych. Parametry tego wielorównaniowego modelu oszacowano na podstawie danych z badania indywidualnych budżetów gospodarstw domowych w Polsce w 2006 roku. Próba roczna obejmowała 37 508 gospodarstw domowych. Uwzględniając oszacowane parametry równań modelu wyznaczono elastyczności cenowe własne i mieszane dla wybranych produktów żywnościowych. Elastyczności cenowe własne dla badanych artykułów żywnościowych były ujemne, najmniej elastycznymi artykułami były jaja (–0,288) a najbardziej elastycznym cenowo – warzywa (–1,118). Mieszane elastyczności cenowe dla badanych artykułów żywnościowych przyjmowały zarówno dodatnie, jak i ujemne wartości, co oznacza, że wśród rozważanych dóbr występowały zależności substytucyjne (dodatnia elastyczność cenowa krzyżowa) oraz komplementarne (ujemna elastyczność cenowa krzyżowa).

63. Beata Bieszk-Stolorz, Iwona Markowicz (Uniwersytet Szczeciński), Analiza szansy wyjścia z bezrobocia w oparciu o przeprowadzoną klasyfikację bezrobotnych według wieku i stażu pracy

Celem artykułu jest analiza szansy wychodzenia z bezrobocia na przykładzie danych z Powiatowego Urzędu Pracy w Szczecinie o osobach wyrejestrowanych w 2009 roku. Jednym z czynników mających wpływ na możliwość uzyskania zatrudnienia jest dotychczasowe doświadczenie zawodowe. Naturalnie, staż pracy jest uzależniony od wieku osoby bezrobotnej. W związku z tym przeprowadzono w pierwszym etapie badań klasyfikację osób poszukujących pracy do grup wieku o zbliżonym ilorazie szans. Drugi etap obejmował badanie szansy podjęcia zatrudnienia w zależności od grupy stażu pracy w ramach poszczególnych grup wieku. Do przeprowadzonych badań wykorzystano model regresji Coxa, umożliwiający wykorzystanie danych cenzurowanych. W artykule poza nowym podejściem do grupowania użyto rzadko stosowane kodowanie zmiennej objaśniającej –1;0;1, które umożliwia porównanie szans analizowanych podgrup z szansą średnią dla grupy.

64. Joanna Olbryś (Politechnika Białostocka), Problem danych niesynchronicznych w modelach market-timing

Problem danych niesynchronicznych nie został do tej pory zbadany w przypadku modeli market-timing polskich funduszy inwestycyjnych. Celem artykułu będzie analiza porównawcza wyników estymacji zarówno klasycznych (H-M oraz T-M), jak też zmodyfikowanych trójczynnikowych (uwzględniających czynniki Famy i Frencha)

modeli market-timing wybranych 15 polskich OFI akcji w okresie styczeń 2003 – czerwiec 2010, w oparciu o dane dzienne, po uwzględnieniu tzw. poprawki Dimsona.

65. Radosław Pietrzyk (Uniwersytet Ekonomiczny we Wrocławiu), Efektywność inwestycji funduszy inwestycyjnych – wykorzystanie modeli market timing

Artykuł jest kontynuacją rozważań autora przedstawionych w artykule „Efektywność inwestycji funduszy inwestycyjnych w okresie hossy i bessy" w którym zostały zaprezentowane modele Treynora-Mazuya oraz Henrikssona-Mertona oraz ich wykorzystanie w ocenie efektywności inwestycji. Celem artykułu jest poszerzenie rozważań o możliwość wykorzystania na polskim rynku modelu Connora-Korajczyka. Model ten pozwala na zniwelowanie wad innych modeli *market timing*, które nie pozwalają na należyte uwzględnienie pozytywnych sygnałów rynku z powodu skośności rozkładu stóp zwrotu z portfela względem stopy zwrotu benchmarku. Skośność ta jest wynikiem wykorzystania przez zarządzających pewnych strategii z wykorzystaniem opcji oraz ubezpieczanie inwestycji. Prezentowany model zakłada wykorzystanie opcji put, której koszt będzie wpływał na wartość wyrazu wolnego równania regresji opisującego model Henrikssona-Mertona. Zaprezentowana metoda zostanie poddana weryfikacji statystycznej, aby ocenić, czy fundusze dostosowują swoje strategie do zmieniających się warunków rynkowych i które z modeli *market timing* są lepiej dopasowane do danych rynkowych.

66. Konstancja Poradowska (Uniwersytet Ekonomiczny we Wrocławiu), Modele dyfuzji innowacji w analizie danych subiektywnych

W przypadku analizy i prognozowania zjawisk nowych, na temat których brak wystarczającej liczby danych empirycznych z przeszłości, budowa formalnego modelu klasycznymi metodami nie jest możliwa. Alternatywnym podejściem może być wykorzystanie w tym celu tzw. danych subiektywnych, czyli opinii pozyskanych od ekspertów merytorycznych, formułowanych na podstawie intuicji i doświadczenia. Spośród formalnych modeli, stosowanych do opisu rozwoju nowych zjawisk najsolidniejsze podstawy teoretyczne zdają się mieć matematyczne modele dyfuzji innowacji. W niniejszym artykule zostanie zaprezentowany sposób konstrukcji prognoz na podstawie takich modeli. Przydatność przedstawione rozwiązań zostanie zweryfikowana w oparciu o materiały z badania *foresigt*: „Zeroemisyjna gospodarka energią w warunkach zrównoważonego rozwoju Polski do 2050".

67. Anna Rutkowska-Ziarko (Uniwersytet Warmińsko-Mazurski w Olsztynie), Alternatywna metoda budowy fundamentalnego portfela papierów wartościowych

Przedmiotem rozważań jest dywersyfikacja ryzyka w portfelu fundamentalnym. W modelu zaproponowanym przez Tarczyńskiego maksymalizowana jest średnia *TMAI* przy ograniczeniach dotyczących między innymi zyskowności oraz ryzyka portfele. W modelu tym rozważa się średnią odchyleń standardowych akcji wchodzących w skład portfela, co nie pokrywa się z odchyleniem standardowym całego portfela. Celem artykułu jest zaproponowanie i empiryczna weryfikacja alternatywnej metody dywersyfikacji ryzyka w portfelu fundamentalnym. Przeprowadzone badania wskazują na możliwość znalezienia portfeli o takiej samej średniej *TMAI* i zyskowności przy

mniejszym ryzyku całkowitym w porównaniu z metodą zaproponowaną przez Tarczyńskiego.

68. Tomasz Stryjewski (Wyższa Szkoła Informatyki i Ekonomii TWP w Olsztynie), Analiza odporności na kryzys przedsiębiorstw z branży budowlanej w zależności od struktury produkcji

Praca prezentuje analizę wyników finansowych grup budowlanych notowanych na GPW w zależności od struktury produkcji. Do badania wyników grup budowlanych zastosowano 4 miary: przychody ze sprzedaży, wynik operacyjny, wynik brutto i netto. Artykuł pozwala zweryfikować hipotezę o zależności pomiędzy strukturą produkcji badanych grup budowlanych a ich odpornością na zawirowania powstałe w 4 kw. 2008 i 1 kw. 2009. Do weryfikacji powyższej hipotezy wykorzystano analizę odległości euklidesowych metodą Warda, taksonomiczne miary rozwoju oraz współczynnik korelacji rang Spearmana.

69. Beata Jackowska, Ewa Wycinka (Uniwersytet Gdański), Modelowanie ryzyka wystąpienia szkody ubezpieczeniowej: budowa i kryteria oceny modelu regresji logistycznej

W artykule podjęto próbę zbudowania modelu ryzyka wystąpienia szkody na polisie ubezpieczeniowej wykorzystując w tym celu regresję logistyczną. Przyjęto hipotezę, że w sytuacji asymetrii informacji na rynku ubezpieczeń ryzyko związane z polisą nie jest addytywne ze względu na ryzyka cząstkowe. Ze względu na specyfikę ryzyka ubezpieczeniowego, w tym małe prawdopodobieństwo wystąpienia szkody, model zbudowano na próbie zbilansowanej. Podczas budowy modelu uwzględniono różne sposoby kodowania zmiennych i ich wpływ na postać modelu. Przy wyborze postaci modelu uwzględniono miary dobroci dopasowania oraz miary zdolności predykcyjnej obliczone dla próby uczącej i testowej. Zaproponowany model pozwala na klasyfikację ubezpieczonych do jednorodnych grup ryzyka na podstawie cech kupowanej polisy.

70. Alina Bojan (Uniwersytet Ekonomiczny we Wrocławiu), Wykorzystanie modeli warunkowej heteroskedastyczności do wyznaczania prawdopodobieństwa niewypłacalności spółek za pomocą ZPP oraz KMV

W artykule scharakteryzowano główne założenia stosowane w stosunkowo nowej i niezbadanej metodzie mierzenia prawdopodobieństwa niewypłacalności spółek opartej na modelu Zero-Price Probability. Do wyznaczenia PD używa on możliwych trajektorii cen akcji opisanych procesem GARCH, a otrzymanych poprzez symulacje Monte Carlo. Na przykładzie wybranych spółek sprawdzono możliwości predykcyjne ZPP, a otrzymane wyniki porównano z rezultatami uzyskanymi z popularnego modelu KMV. Pozwoliło to odpowiedzieć na dwa zasadnicze pytania: czy nowa metodologia zaproponowana przez Fantazziniego prowadzi do dokładniejszych oszacowań niż KMV i czy w ogóle może być ona przydatnym narzędziem do pomiaru ryzyka na rynku polskim.

71. Witold Hantke (Wyższa Szkoła Zarządzania i Nauk Społecznych im. ks. Emila Szramka w Tychach), Klasyfikacja branżowych rynków pracy województwa śląskiego w świetle integracji europejskiej i kryzysu ekonomicznego

Na śląski rynek pracy w ostatnich latach istotny wpływ wywarły dwa czynniki: integracja europejska i ogólnoświatowy kryzys gospodarczy. Niniejszy artykuł częściowo odpowiada na pytanie, czy wpływ ten był równomierny w odniesieniu do wszystkich środowisk zawodowych. W pierwszej jego części za pomocą metod taksonomicznych i dyskryminacyjnych grupy zawodowe zostały podzielone na jednorodne klasy. Następnie, przy użyciu technik porządkowania liniowego, został utworzony ranking pozwalający wyróżnić te zawody, które skorzystały na akcesji Polski do UE w największym stopniu oraz te, które odczuły najsilniej skutki kryzysu gospodarczego.

72. Monika Książek (Szkoła Główna Handlowa), Karolina Keler (Uniwersytet Jagielloński), Analiza oszczędności i kredytów polskich gospodarstw domowych z użyciem modelu ukrytego łańcucha Markowa

W artykule opracowano typologię stanów oszczędności i kredytów polskich gospodarstw domowych z wykorzystaniem modelu ukrytego łańcucha Markowa. Charakterystykę stanów oparto na posiadaniu, wysokości i przeznaczeniu oszczędności i zobowiązań oraz formie oszczędności i źródle kredytu. Wyróżniono i scharakteryzowano 11 stanów, oszacowano ich trwałość oraz wskazano najczęstsze kierunki przejść pomiędzy nimi. Uzyskane wyniki świadczą o występowaniu zróżnicowanych i trwałych postaw wobec gospodarowania własnymi pieniędzmi.

73. Michał Kukliński (Uniwersytet Mikołaja Kopernika w Toruniu), Studium porównawcze metod analizy koszykowej

Artykuł przedstawia porównanie metod analizy koszykowej na przykładzie transakcyjnej bazy danych. W publikacji przedstawione zostały poszczególne etapy przygotowania danych oraz analizy za pomocą oprogramowania STATISTICA oraz SPSS Clementine. Zestawienie podstawowych charakterystyk metod *a'priori* oraz *GRI* pozwala na wybór odpowiedniego algorytmu, w zależności od typu danych oraz ilości danych wejściowych.

74. Monika Osińska (Uniwersytet Ekonomiczny w Poznaniu), Mierniki oceny jakości podziału – analiza porównawcza

Ocena jakości podziału jest jednym z czterech zasadniczych etapów analizy skupień istotnie wpływającym na interpretację uzyskanych wyników. W literaturze przedmiotu istnieje duża liczba wskaźników wspomagających proces wyboru podziału optymalnego, jednak spora ich część przejawia pewne własności, które w dużej mierze mogą ograniczać obszary ich zastosowań. Głównym celem referatu jest zaproponowanie nowego wskaźnika oceny jakości grupowania – wskaźnika *CNI* oraz porównanie jego użyteczności z siedmioma najbardziej znanymi w literaturze wskaźnikami. Obiekty poddane analizie opisane zostały przy użyciu zmiennych dwuwymiarowych, natomiast kolejne podziały uzyskano w wyniku zaimplementowania metody *k*-średnich. Wszystkie obliczenia wykonano w programie *R*.

75. Małgorzata Śniegocka-Łusiewicz (Uniwersytet Mikołaja Kopernika w Toruniu), Analiza koszykowa w badaniu cykliczności reguł asocjacji w handlu detalicznym

W pracy zakłada się istnienie zmienności w czasie w preferencjach konsumentów odnośnie wyboru produktów. Wiedza na temat reguł w wyborze następników oraz praw-

dopodobieństwie wyboru poprzedników w poszczególne dni oraz miesiące umożliwi lepsze przygotowanie i skierowanie reklamy do reklamobiorców. Analiza koszykowa umożliwia przebadanie empirycznego zbioru danych i uwidacznia zmiany w poziomie ufności i wsparcia w poszczególnych okresach. Badanie uwidoczniło istnienie najsilniejszych reguł w miesiącu październiku oraz w soboty, przy równoczesnym istnieniu największej liczby transakcji. Ponadto wykazane zostało, że w badaniach empirycznych kryteria minimalne analizy koszykowej muszą być na niskim poziomie.

76. Mirosław Krzyśko, Karol Deręgowski (Uniwersytet im. Adama Mickiewicza w Poznaniu), Nieliniowe składowe główne

Hotelling (1933) zaproponował liniową metodę redukcji wymiarowości zwaną analizą składowych głównych. Następnie Scholkopf, Smola i Muller (1998) podali konstrukcje nieliniowych składowych głównych za pomocą funkcji jądrowych. W opracowaniu tym pokazujemy nowa metodę konstrukcji nieliniowych składowych głównych i jej związek z metodą Scholkofa, Smoli i Mullera (1998).

77. Krzysztof Kompa (Szkoła Główna Gospodarstwa Wiejskiego w Warszawie), Porównanie efektywności giełd z rynków wschodzących za pomocą metod porządkowania liniowego

Proces transformacji krajów tzw. Bloku Wschodniego trwa juz 20 lat i dotyczy wszystkich sektorów gospodarki, w tym sektora finansowego. Naturalnym wiec staje sie pytanie na ile wprowadzone przemiany pozwoliły przybliżyć sie tym krajom do poziomu rynków rozwiniętych. Zazwyczaj prowadzone analizy dotyczą giełd papierów wartościowych utożsamianych z rynkiem kapitałowym. W prowadzonych badaniach rozważane są giełdy, będące oczywiście instytucjami rynku kapitałowego, ale jako przedsiębiorstwa o różnej formie własności, ale realizujący podstawowy cel jakim jest osiąganie jak najlepszych wyników finansowych. Celem prowadzonych analiz jest zbadanie efektywności 23 giełd europejskich zrzeszonych w Federacji Giełd Europejskich FESE. W badaniach uwzględniono dane dotyczące wyników finansowych instytucji rynku kapitałowego, traktowanych jako przedsiębiorstwa. Baza danych obejmuje dane roczne za lata 2004-2008. Zastosowano miary porządkowania liniowego, skonstruowane dla różnych zestawów zmiennych diagnostycznych oraz wag. Na ich podstawie przeprowadzono grupowanie obiektów i zbadano stabilność mierników oraz klasyfikacji na wprowadzane zmiany parametrów oraz metod normalizacji zmiennych.

78. Tomasz Górecki, Mirosław Krzyśko (Uniwersytet im. Adama Mickiewicza w Poznaniu), Regresyjne metody łączenia klasyfikatorów

LeBlanc i Tibshirani (1996) zaproponowali tak zwaną regresję stosową jako metodę łączenia informacji płynących z różnych klasyfikatorów, zamiast poszukiwania najlepszego klasyfikatora w pewnym zbiorze klasyfikatorów. Klasyfikator połączony jest liniową kombinacją estymatorów prawdopodobieństw a posteriori otrzymanych z poszczególnych klasyfikatorów. W opracowaniu porównano metodę regresji stosowej z innymi metodami regresyjnymi, takimi jak: regresja logistyczna, regresja grzbietowa, regresja składowych głównych, regresja częściowych najmniejszych kwadratów, nieparametryczna regresja jądrowa, regresja LASSO, LARS oraz LARSEN.

79. Daniel Papla (Uniwersytet Ekonomiczny we Wrocławiu), Modelowanie zarażania rynków finansowych – teoria i badania empiryczne

Dotychczas brak jest ogólnie przyjętej definicji zarażania rynków finansowych (contagion in financial markets), najczęściej jednak przyjmuje sie, że zarażanie występuje wtedy, kiedy podczas kryzysu obserwuje sie znacząco zwiększoną zależność między ruchami cen na różnych rynkach finansowych. Kryzysy finansowe są ważnym zjawiskiem dla całej gospodarki, ponieważ podczas kryzysu zwiększają sie koszty pośrednictwa oraz koszty kredytów, trudniejszy jest również dostęp do kredytów. Powoduje to ograniczenia działalności sektora realnego, co może prowadzić do kryzysu również i w tym sektorze. Dosyć duża częstość występowania kryzysów finansowych może prowadzić do wniosku, ze sektor finansowy jest szczególnie wrażliwy na różnego rodzaju zaburzenia. Jedna z teorii, która to tłumaczy jest właśnie teoria zarażania rynków finansowych. W tej teorii małe zaburzenia, które początkowo dotyczą jedynie kilku instytucji lub wybranego regionu, rozprzestrzeniają sie jak epidemia na cały sektor finansowy, a potem na całą gospodarkę. Aby móc dokładniej przeanalizować zarażanie rynków finansowych, w niniejszym artykule przedstawione zostaną wybrane modele tego zjawiska. Zaprezentowane również będą badania empiryczne służące weryfikacji tych modeli.

80. Joanna Urban (Politechnika Białostocka), Klasyfikacja polskich jednostek naukowych w ocenie parametrycznej MNiSzW

Ewaluacja wyników dzialalnosci naukowej jest waznym instrumentem zarzadzania organizacja i finansowaniem badan naukowych. Stanowi istotny element realizacji polityki naukowej i ukierunkowywania dzialalnosci jednostek sektora naukowego. Ze wzgledu na zlozonosc dzialalnosci naukowej, ocena jednostek naukowych na poziomie krajowym jest wyjatkowo trudna i skomplikowana. W Polsce, powszechnie obowiazujacym systemem oceny jednostek naukowych ubiegajacych sie o dotacje z budzetu panstwa jest ocena parametryczna Ministerstwa Nauki i Szkolnictwa Wyzszego. Ocena parametryczna jest podstawa ustalania kategorii jednostek naukowych oraz przyznawania srodków finansowych na prowadzenie dzialalnosci naukowej. Ocena parametryczna jednostek naukowych jest szeroko dyskutowana w srodowisku co do zasad, uzytecznosci oraz kryteriów oceny. Przeprowadzana jest na podstawie zestawu zmiennych zgrupowanych w filary. Kazdemu z filarów nadano wagi, z uwzglednieniem specyfiki tzw. grup jednorodnych. Dyskusyjne zagadnienie stanowi klasyfikacja jednostek naukowych w grupy jednorodne i ich rozlacznosc. W artykule podjeto zagadnienie separowalnosci grup jednorodnych. Zbadano zróznicowanie jednostek ze wzgledu na przyjete kryteria oceny oraz jaka sile róznicujaca jednostki naukowe wewnatrz grup maja poszczególne zmienne. Do ilustracji tego problemu zastosowano metode graficznej analizy danych wielowymiarowych (RGM), która umozliwia wspólne przedstawienie na plaszczyznie relacji obiekt-obiekt, cecha-cecha oraz obiekt-cecha. Analizy przeprowadzono na podstawie danych bedacych podstawa do oceny parametrycznej w roku 2006 i wskazano rekomendacje klasyfikacji polskich jednostek naukowych w grupy jednorodne.

81. Marcin Owczarczuk (Szkoła Główna Handlowa w Warszawie), Prognozowanie nadużyć na rynku telefonii komórkowej

W opracowaniu przedstawiono poważny problem, z jakim stykają sie operatorzy telefonii komórkowej, mianowicie nadużycia związane z wyłudzeniem drogich modeli aparatów telefonicznych sprzedawanych klientom po promocyjnych cenach przy podpisywaniu umowy abonenckiej. Celem operatora jest w takiej sytuacji prognozowanie, którzy klienci są skłonni do dokonania tego typu nadużyć i np. ustalenie dla nich odpowiedniej kaucji. Zmienna objaśniana jest binarna i oznacza fakt opłacenia chociaż jednego rachunku za telefon, natomiast zmienne objaśniające są konstruowane na podstawie wniosku aplikacyjnego i dotyczą takich informacji, jak wysokość abonamentu czy dokument, którym się posłużył klient. W badaniu porównano szereg 36 klasyfikatorów zbudowanych w oparciu o dane pochodzące od jednego z polskich operatorów sieci komórkowej. W szczególności do porównań wykorzystano metody dające modele interpretowalne takie jak regresja logistyczna czy drzewa klasyfikacyjne oraz techniki pozbawione tej własności jak metoda wektorów nośnych czy sieci neuronowe.

82. Ewa Nowakowska (Szkoła Główna Handlowa w Warszawie), W poszukiwaniu skupień – nowe podejście do oceny wyrazistości struktury danych

Praca poświecona jest zagadnieniu oceny wyrazistości struktury skupień (ang. clusterability), czyli tak zwanej klastrowalności. Jest to problem słabo rozpoznany teoretycznie, aczkolwiek o dużym znaczeniu praktycznym. Możliwość wstępnej oceny szans uzyskania stabilnego podziału segmentacyjnego przed przystąpieniem do analiz pozwoliłaby zoptymalizować prace i wysiłek, dobrać adekwatną przestrzeń zmiennych aktywnych, a nawet oszacować nieznane parametry (np. liczbę skupień). Celem pracy jest prezentacja nowego podejścia do problemu, osadzonego w kontekście statystycznej analizy danych i opartego na modelu probabilistycznym. Jako model probabilistyczny przyjęta zostaje heteroskedastyczna mieszanka rozkładów normalnych. W tym kontekście, w języku rozkładu generującego dane, zdefiniowane zostaje pojecie klastrowalności jako prawdopodobieństwo poprawnej klasyfikacji. Jest ono wyznaczane jako 1 - p, gdzie $p$to wielkość obszaru nakładania sie rozkładów tworzących mieszankę. Jest to miara ujmująca kluczowe parametry modelu (wzajemne położenie środków rozkładów oraz rozproszenie) decydujące o późniejszej klastrowalności wygenerowanych danych. Jednocześnie jako skalar z przedziału [0,1] nie nastręcza trudności interpretacyjnych. W drugim kroku ta teoretyczna miara przełożona zostaje na praktyczny język estymacji, co umożliwia szacowanie jej wartości dla konkretnych zbiorów danych. Aby ograniczyć złożoność obliczeniową, estymacja oparta jest na losowym przeszukiwaniu przestrzeni danych, natomiast w procesie analizy wykorzystywane jest rzutowanie znaczących podzbiorów na podprzestrzenie jednowymiarowe. Analiza jednowymiarowych rozkładów prowadzi do oszacowania położenia obszarów o relatywnie dużym zagęszczeniu obserwacji oraz rozproszenia wokół ich środków, co w efekcie pozwala oszacować wyrazistość struktury skupień bez wcześniejszej znajomości rozwiązania segmentacyjnego.

83. Jan Żółtowski, Iwona Staniec, Jacek Stańdo (Politechnika Łódzka), Identyfikacja wpływu cech społecznych i ekonomicznych na wyniki nauczania

Znajomość wpływu cech społecznych i ekonomicznych na wyniki nauczania pozwoli na odpowiednie sterowanie procesem nauczania, szczególnie w początkowych latach nauki. Coraz więcej pojawia się prac naukowych poświeconych nie tylko samej szkole i jej procesowi wewnętrznemu (środowisku), ale skupiających sie na identyfikowaniu wpływu na proces nauczania warunków zewnętrznych, takich jak czynniki społeczne czy ekonomiczne. Celem przedstawionych rozważań jest klasyfikacja cech społecznych i ekonomicznych ze względu na wyniki próbnych egzaminów gimnazjalnych i maturalnych. Dzięki niej możliwe jest wskazanie cech charakteryzujących poszczególne grupy uczniów oraz wyjaśnienie, jakimi czynnikami różnią się wyodrębnione klasy. Podstawą interpretacji są zmienne biorące udział w procesie klasyfikacji. Badaniem objęto ok. 3000 uczniów zdających egzamin gimnazjalny i maturalny w całym kraju. W procesie identyfikacji wykorzystano wielokryterialne metody porządkowania liniowego, wzorzec rozwoju Hellwiga, analizę skupień oraz regresję logitową.

W pierwszym dniu konferencji odbyło się posiedzenie członków Sekcji Klasyfikacji i Analizy Danych Polskiego Towarzystwa Statystycznego, któremu, po wyborze, przewodniczył prof. UMK dr hab. Tadeusz Kufel. Ustalono plan przebiegu zebrania obejmujący następujące punkty:

1. Sprawy różne.
2. Sprawozdanie z działalności Sekcji Klasyfikacji i Analizy Danych PTS.
3. Informacje dotyczące konferencji zagranicznych.
4. Zapowiedzi kolejnych konferencji SKAD PTS.
5. Uchwała w sprawie powołania członków honorowych Sekcji SKAD PTS.
6. Wybór Rady Sekcji SKAD PTS na kadencję 2011-2012.

Na wstępie prof. dr hab. Marek Walesiak poinformował zebranych, że od 8 marca 2010 r. obowiązuje (zamieszczony na stronie www.skad.pl) nowy Regulamin Sekcji Klasyfikacji i Analizy Danych (SKAD) Polskiego Towarzystwa Statystycznego (PTS) zatwierdzony na posiedzeniu Rady Głównej Polskiego Towarzystwa Statystycznego, które odbyło się w dniu 8 marca 2010 r. w Krakowie oraz nowa Deklaracja Członka Sekcji Klasyfikacji i Analizy Danych (SKAD) Polskiego Towarzystwa Statystycznego.

Sprawozdanie z działalności Sekcji Klasyfikacji i Analizy Danych w okresie wrzesień 2009 – wrzesień 2010 przedstawił przewodniczący Rady Sekcji prof. dr hab. Marek Walesiak. Na początku poinformował zebranych, że:

– SKAD liczy 217 członków,

– opublikowany został zeszyt nr 17 z serii *Taksonomia* pt. „*Klasyfikacja i analiza danych – teoria i zastosowania*" z konferencji Sekcji Klasyfikacji i Analizy Danych PTS, która odbyła się w Międzyzdrojach w dniach 16-18.09.2009 r.,

– w numerze nr 2/2010 *Przeglądu Statystycznego* ukaże się sprawozdanie z konferencji z 2009 r.

W kolejnej części zebrania przekazano informacje dotyczące konferencji zagranicznych. Prof. Marek Walesiak przekazał zaproszenie Niemieckiego Towarzystwa Klasy-

fikacyjnego na 35 Konferencję GfKl, która odbędzie się w 2011 roku we Frankfurcie nad Menem. Prof. Krzysztof Jajuga poinformował, że konferencja IFCS, która miała mieć miejsce w 2011 roku w St. Andrews nie odbędzie się. Przyczyną odwołania konferencji było niedotrzymanie wymogów formalnych przez organizatorów. Przekazał również, że zorganizowano inną konferencję o bardzo zbliżonym zakresie tematycznym. Prof. UEK dr hab. Andrzej Sokołowski dodał, że konferencję nazwano „Not IFCS Conference", a obecnie figuruje ona pod nazwą International Classification Conference (ICC). W komitecie organizacyjnym konferencji ICC umieszczono, bez uzyskania zgody zainteresowanych, nazwiska osób będących w komitecie konferencji IFCS.

W następnym punkcie posiedzenia podjęto kwestię organizacji kolejnych konferencji SKAD. Prof. UP dr hab. Feliks Wysocki z Uniwersytetu Przyrodniczego w Poznaniu potwierdził wolę organizacji przyszłorocznej konferencji SKAD we współpracy z Uniwersytetem Ekonomicznym w Poznaniu. Poinformował, że konferencja odbędzie się prawdopodobnie w Wągrowcu w dniach 21-23 września 2011 roku w hotelu Pietrak. Propozycja ta została przyjęta przez aklamację. Zasugerowano, aby opłata konferencyjna nie uległa zmianie i wynosiła 1.200 zł, a dla doktoranta 1.000 zł.

Prof. dr hab. Eugeniusz Gatnar zrezygnował z organizacji konferencji SKAD w 2012 roku. Z kolei prof. dr hab. Joanicjusz Nazarko poinformował zebranych, że podejmie się zorganizowania konferencji w 2012 roku.

W dalszej części zebrania prof. dr hab. Krzysztof Jajuga przedłożył wniosek o powołanie członków honorowych Sekcji SKAD, szczególnie zasłużonych dla Sekcji, w osobach prof. dr. hab. Kazimierza Zająca i prof. dr. hab. Zdzisława Hellwiga (współtwórców Sekcji) oraz prof. dr hab. Stanisławy Bartosiewicz.

Przewodniczący zebrania Prof. Tadeusz Kufel powołał Komisję Skrutacyjną w następującym składzie: dr Joanna Landmesser, dr Jacek Batóg i dr Tomasz Stryjewski. Odbyło się głosowanie, a następnie Komisja ogłosiła wyniki:

— prof. dr hab. Zdzisław Hellwig: 64 głosy za, 0 osób się wstrzymało, 0 głosów przeciw,
— prof. dr hab. Kazimierz Zając: 64 głosy za, 0 osób się wstrzymało, 0 głosów przeciw,
— prof. dr hab. Stanisława Bartosiewicz: 63 głosy za, 1 osoba się wstrzymała, 0 głosów przeciw.

Uchwała została przyjęta.

W kolejnej części zebrania dokonano wyboru członków Rady Sekcji SKAD PTS na kadencję w okresie 2011-2012. Prof. Tadeusz Kufel przypomniał skład Rady obecnej kadencji. Prof. Krzysztof Jajuga zaproponował zmniejszenie liczby członków Rady z 8 do 5. Komisja Skrutacyjna przeprowadziła głosowanie jawne i przekazała następujące wyniki: 59 głosów za, 3 osoby się wstrzymały, 2 głosy przeciw. Wniosek o zmniejszenie liczby członków Rady SKAD z 8 do 5 osób został przyjęty.

Następnie prof. Krzysztof Jajuga złożył propozycję kandydatów do Rady w osobach prof. Marka Walesiaka, prof. Andrzeja Sokołowskiego, prof. Eugeniusza Gat-

nara i dra Krzysztofa Najmana. Prof. Marek Walesiak zaproponował kandydaturę prof. Krzysztofa Jajugi. Nie było dalszych zgłoszeń.

Odbyło się głosowanie tajne. Komisja podliczyła głosy:

— prof. Marek Walesiak: Tak – 64 głosy, Nie – 0 głosów, nieważne 2 głosy,
— prof. Andrzej Sokołowski: Tak – 64 głosy, Nie – 1 głos, nieważny 1 głos,
— prof. Eugeniusz Gatnar: Tak – 63 głosy, Nie – 2 głosy, nieważny 1 głos,
— prof. Krzysztof Jajuga: Tak – 64 głosy, Nie – 0 głosów, nieważne 2 głosy,
— dr Krzysztof Najman: Tak – 64 głosy, Nie – 0 głosów, nieważne 2 głosy.

Tym samym wszyscy zgłoszeni kandydaci zostali wybrani do Rady Sekcji SKAD PTS. Następnie poinformowano zgromadzonych, iż Rada dokona swojego ukonstytuowania po zebraniu, zaś wyniki zostaną przekazane członkom Sekcji w czasie uroczystej kolacji.

W trakcie uroczystej kolacji prof. Tadeusz Kufel poinformował uczestników konferencji, iż Rada Sekcji się ukonstytuowała i nowym przewodniczącym na kadencję 2011-2012 został prof. dr hab. Marek Walesiak, zastępcą przewodniczącego – prof. UEk dr hab. Andrzej Sokołowski, sekretarzem zaś – dr Krzysztof Najman.

PRZEGLĄD STATYSTYCZNY
R. LVIII – ZESZYT 1-2 – 2011

RECENZJE

JANUSZ L. WYWIAŁ

WPROWADZENIE DO METODY REPREZENTACYJNEJ,

WYDAWNICTWO AKADEMII EKONOMICZNEJ W KATOWICACH, KATOWICE 2010, S. 132.

Metoda reprezentacyjna znajduje szerokie zastosowania w wielu badaniach statystycznych prowadzonych przez urzędy statystyczne różnych krajów, w badaniach opinii publicznej, w badaniach ekonomicznych i społecznych oraz różnych badaniach naukowych. Metoda ta wykładana jest na niektórych wyższych uczelniach, a także prowadzone są szkolenia personelu statystycznego urzędów statystycznych z podstaw tej metody. Kontynuowane są prace naukowo-badawcze nad rozwojem teorii metody reprezentacyjnej i jej zastosowań w praktyce.

W Polsce metoda reprezentacyjna stosowana jest od wielu lat w badaniach statystycznych Głównego Urzędu Statystycznego, a także w wielu badaniach społeczno-ekonomicznych realizowanych przez inne instytucje i ośrodki naukowe. Właśnie w Polsce fundamentalny wkład w rozwój teorii metody reprezentacyjnej wniósł polski statystyk, ***prof. Jerzy Neyman,*** który przedstawił najpierw swoje propozycje w języku polskim w 1933 r. w pracy pt. „*Zarys teorii i praktyki badania struktury ludności metodą reprezentacyjną",* a w języku angielskim w słynnym artykule z 1934 r.[1] który zmienił teoretyczne podstawy wnioskowania na podstawie badań częściowych, wprowadzając błędy losowe oparte na randomizacji rozkładu[1]. Wprowadził *optymalną lokalizację próby* w losowaniu warstwowym, a także *przedziały ufności*[2] Użył on terminu „*reprezentatywny"* w nowym znaczeniu. Neyman zalecał podejście probabilistyczne

[1] J. Neyman (1934): On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection. *Journal of the Royal Statistical Society*, s.558-606.

[2] J. Neyman (1935): On the Problem of Confidence Intervals, *The Annals of Mathematical Statistics,* J. Neyman (1938), Contribution to the Theory of Sampling Human Population, *Journal of American Statistical Association*, nr 33, s.101-116.

(randomizacyjne) przy wyborze próby, a krytycznie odnosił się do jej celowego wyboru. Podejście J. Neymana zwyciężyło i jest powszechnie stosowane w praktyce badań niewyczerpujacych.

Metoda reprezentacyjna jest od dawna w Polsce wykładana w wielu szkołach wyższych, twórczo rozwijana i stosowana w praktyce. W 1962 r. został opublikowany pierwszy podręcznik z metody reprezentacyjnej przygotowany przez R. Zasępę[3], a w następnych latach zostały opublikowane dalsze książki poświęcone tej metodzie badawczej[4].

W Polsce organizowane są systematycznie międzynarodowe konferencje poświęcone pracom badawczym nad zastosowaniem metody reprezentacyjnej w badaniach ekonomicznych i społecznych, których organizatorem jest prof. dr hab. Janusz L. Wywiał z Uniwersytetu Ekonomicznego w Katowicach. Prof. J. L. Wywiał od wielu lat prowadzi intensywne badania naukowe nad teorią metody reprezentacyjnej, a ostatnio wydał podręcznik pt. „***Wprowadzenie do metody reprezentacyjnej”***.

Uważam, że publikacja ta wypełnia lukę na naszym rynku wydawniczym w zakresie wprowadzenia do ogólnej teorii metody reprezentacyjnej i na pewno przyczyni się do głębszego zrozumienia podstaw teorii metody reprezentacyjnej.

Publikacja ta składa się ze wstępu, trzech rozdziałów, indeksu oraz bibliografii. Autor we wstępie zaznacza, że, praca ta ma dać czytelnikowi możliwie prosty obraz zagadnień, jakimi zajmuje się metoda reprezentacyjna. Dlatego też dużo miejsca poświęca na proste, lecz dość szczegółowe przedstawienie podstaw mechanizmów uogólnienia wyników przetwarzania danych obserwowanych w próbie na własności całej populacji. Obok znanego randomizacyjnego podejścia w metodzie reprezentacyjnej, wzięto również pod uwagę problem wnioskowania modelowego (predykcyjnego), jako że jest to w wielu sytuacjach praktycznych racjonalne podejście, pozwalające pominąć – zdaniem Autora – klasyczny etap losowania próby, który jest zastępowany celowym wyborem próby na podstawie racjonalnych przesłanek wynikających z już dostępnej informacji o populacji. W tym przypadku – jak twierdzi Autor – umiejętne korzystanie z dodatkowych, dostępnych informacji o populacji prowadzi do wartościowych prób umożliwiających wnioskowanie statystyczne o parametrach populacji. W szczególności dotyczy to wnioskowania na podstawie danych pochodzących z rejestrów administracyjnych. Mam jednak poważne wątpliwości, czy korzystanie „*z dodatkowych, dostępnych informacji o populacji prowadzi do wartościowych prób umożliwiających dokładne wnioskowanie statystyczne o parametrach populacji”*. Nie mamy przecież informacji o dokładności danych dodatkowych, a wiadomo, że rejestry administracyjne obarczo-

[3] Zasępa R. (1962): *Badania statystyczne metodą reprezentacyjną*, PWN, Warszawa.

[4] Zasępa, R. (1972), *Metoda reprezentacyjna*, PWE, Warszawa.
Pawłowski Z. (1972). *Wstęp do statystycznej metody reprezentacyjnej*, PWN, Warszawa,
Steczkowski J. (1988). Zastosowanie metody reprezentacyjnej w badaniach społeczno-ekonomicznych, PWN, Warszawa,
Bracha Cz. (1996), *Teoretyczne podstawy metody reprezentacyjnej,* PWN, Warszawa,
Szreder M. (2010*): Metody i techniki sondażowych badań opinii*. PWE, Warszawa.

ne są różnego rodzaju błędami. Dodatkowe informacje na pewno mogą zwiększyć precyzję ocen, uzyskanych z badań reprezentacyjnych, ale nie zapewniają dokładnego wnioskowania o parametrach populacji. Odniosłem wrażenie, że Autor akceptuje *„celowy wybór próby na podstawie racjonalnych przesłanek wynikających z już dostępnej informacji o populacji"*. Przecież wnioskowanie modelowe (predykcyjne) opiera się na szeregu założeniach, które nie zawsze są spełnione w rzeczywistości, a ponadto błędy w danych dodatkowych uniemożliwiają uzyskanie dokładnych ocen. Szkoda, że Autor nie wprowadził w swojej pracy pojęcia precyzji ocen, aby wyjaśnić czytelnikowi, iż oceny z badań reprezentacyjnych mogą być jednocześnie precyzyjne, a zarazem niedokładne. To zagadnienie warto wyjaśnić do końca, aby zrozumieć istotę badań reprezentacyjnych.

Pracę tak skonstruowano, aby przedstawić możliwości oceny podstawowych parametrów w rożnych sytuacjach za pomocą narzędzi dostarczanych zarówno przez podejście randomizacyjne, jak i modelowe wnioskowania statystycznego.

W rozdziale 1, pt. ***Próba losowa i statystyki***, przedstawiono populację i parametry jej cech. Przede wszystkim zdefiniowano i omówiono podstawowe pojęcia. Podano definicję próby oraz jej planu i schematu losowania, a także szerzej omówiono wybrane plany losowania prób nieprostych. Znajduje się tu wprowadzenie w podstawy konstrukcji planów i schematów losowania próby, przedstawiono też własności wybranych planów i schematów. Dalej zdefiniowano *dane* i *statystyki,* ilustrując je szeroko przykładami. Rozdział ten zawiera wiadomości, które są wykorzystane w dalszej treści pracy.

Rozdział 2 dotyczy ***estymacji.*** Omówiono w nim podejście randomizacyjne oraz zdefiniowano estymator badanego parametru, a następnie wyjaśniono podstawowe własności estymatorów wybranych parametrów, podając proste przykłady wyprowadzania ich rozkładów prawdopodobieństwa. Tak jak w podręczniku Lohr (1999)[5], wyjaśniono podstawy wnioskowania randomizacyjnego, jak i modelowego, przechodząc stopniowo od prostych zagadnień do bardziej złożonych problemów, w których brane są pod uwagę m.in. estymatory lub predykatory ilorazowe i regresyjne. I tak, zdefiniowano:

- estymator $\boldsymbol{t}_S$parametru θ,
- błąd estymacji: $u(t_S) = t_s - \theta$,
- obciążenie estymatora: $b(t_S) = E[u(t_s)] - \theta$,
- względny wskaźnik obciążenia: $wb(t_S) = \dfrac{E(t_S) - \theta}{|\theta|} 100\%$,
- błąd średniokwadratowy estymacji: $MSE(t_S) = D^2(t_S) + b^2(t_S)$.

W metodzie reprezentacyjnej przyjmuje się, że błąd średniokwadratowy estymacji, a szczególnie jego pierwiastek kwadratowy, jest miarą dokładności oceny (s. 36-37). Niestety, w wielu badaniach reprezentacyjnych, a szczególnie w badaniach społecznych, w których zwykle uzyskuje się oceny obciążone, na skutek różnych przyczyn

[5] Lohr S.L. (1999): *Sampling: Design and Analysis.* Duxbury Press, International Thompson Publishing Company.

(np. odmowy udziału w badaniach, błędy odpowiedzi i pomiaru), nie powinno stosować się oceny pierwiastka błędu średniokwadratowego, gdy nie podjęto dodatkowo oceny obciążenia. W statystyce matematycznej zwykle za $b(t_S)$ przyjmuje się obciążenie estymatora t_S , ale w metodzie reprezentacyjnej $MSE(t_S)$ jest średnią miarą dokładności oceny. Jeśli nie podjęto próby oceny obciążenia, wtedy należy ograniczyć się tylko do podania oceny precyzji, tj. $D(t_S)$ lub względnej precyzji, tj. $CV(t_S) = \frac{D(t_S)}{\hat{\theta}}$. Często w badaniach reprezentacyjnych uzyskuje się oceny bardzo precyzyjne a zarazem niedokładne. Dotyczy to wielu ocen uzyskanych metodą wywiadu, np. z zakresu dochodów, w których obciążenie nieraz przewyższa 10%, a precyzja kształtuje się poniżej 2%. W takich przypadkach nie ma sensu budowania przedziałów ufności, gdyż wprowadza się użytkownika w błąd, twierdząc, że *„podany przedział ufności pokrywa prawdziwą wartość szacowanego parametru przy danym poziomie ufności*". (s. 39). Warto o tym pamiętać, przy planowaniu badania reprezentacyjnego, a następnie przy analizie uzyskanych wyników. Z tych powodów prof. J. Neyman nie zalecał nam stosowania metody reprezentacyjnej w badaniach budżetów rodzinnych, w 1958 r. podczas 6-tygoniowej konsultacji dla Komisji Matematycznej GUS[6]. A przecież we wszystkich badaniach reprezentacyjnych prowadzonych przez urzędy statystyczne na wielką skalę pewna frakcja badanych jednostek nie podejmuje badań, występują błędy odpowiedzi lub pomiaru oraz innego rodzaju błędy nielosowe. W takich przypadkach nie możemy twierdzić, iż podany przedział ufności pokrywa prawdziwą wartość parametru na określonym poziomie ufności. Jesteśmy w stanie tylko określić wielkość precyzji.

Powszechnie stosowane w praktyce procedury wnioskowania przedstawiono możliwie szczegółowo. Dotyczy to estymacji lub predykcji wartości globalnej bądź jej średniej na podstawie prób warstwowych, grupowych, dwustopniowych. Zasygnalizowano również problem optymalnego ustalania liczebności prób, które maja być losowane np., z warstw, na jakie wcześniej podzielono populację.

Problem racjonalności ustalania liczebności losowanych prób jest jednym z podstawowych problemów planowania badania reprezentacyjnego i bezpośrednio rzutuje na koszty prowadzenia takiego badania. Możliwości wykorzystania prezentowanych metod wnioskowania ilustrowano licznymi przykładami.

Szeroko potraktowano strategie estymacji zależne od cechy dodatkowej. Przedstawiono estymatory i predykatory ilorazowe oraz regresyjne, a także weryfikacje regresyjnego modelu nadpopulacji. Rozdział ten kończy syntetyczne przedstawienie wybranych metod oceny wariancji estymatorów złożonych: *linearyzację Taylora*, *technikę jackknife'a* oraz *metodę bootstrapa.* Metody te były i dalej są stosowane w złożonych badaniach reprezentacyjnych w Polsce i są szeroko opisane w innych publikacjach[7].

[6] Zasępa, R. (1958), Problematyka badań reprezentacyjnych GUS w świetle konsultacji z prof. J. Neymanem, *Wiadomości Statystyczne,* nr 6, s. 7-12.

[7] J. Kordos, A. Zięba-Pietrzak (2010), Development of standard error estimation methods in complex household sample surveys in Poland, Statistics in Transition – new series, Vol. 11, No. 2.

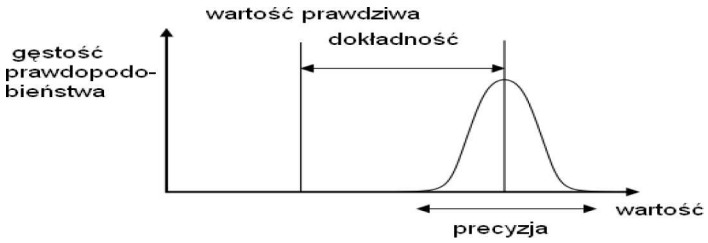
W rozdziale 3 przedstawiono ***popularne strategie estymacji***. Rozpoczęto od średniej z próby systematycznej, rozważając prostą próbę systematyczną, próbę z prawdopodobieństwem inkluzji proporcjonalnej do wartości zmiennej dodatkowej oraz wnioskowanie przy modelu populacji uporządkowanej. Dalej przedstawiona jest średnia z próby warstwowej, która jest najczęściej stosowana w praktyce. Tutaj Autor stwierdza, że „*wariancja* $D^2_S(\bar{y}_{S_{hwS}}, P_w)$ *służy do oceny precyzji estymacji wartości globalnej cechy w h-tej warstwie*". Wydaje się, że warto tu wyjaśnić różnicę między pojęciami precyzji i dokładności. Na s. 36 podano dekompozycję błędu średniokwadratowego estymacji:

$$MSE(t_S) = D^2(t_S) + b^2(t_S)$$

gdzie $D^2(t_S)$ jest wariancją estymatora. Pierwiastek z błędu średniokwadratowego $MSE(t_S)$ jest nazywany błędem średnim szacunku parametru θ, który określa przeciętny błąd oceny parametru θ za pomocą estymatora $\boldsymbol{t_S}$**.** Pierwiastek z wariancji estymatora jest odchyleniem standardowym $D(t_S)$estymatora, określającym przeciętne odchylenie wartości estymatora od jego obserwowanej wartości oczekiwanej. Właśnie ***parametr*** $\boldsymbol{D(t_S)}$ ***reprezentuje precyzję estymatora***, tj. przeciętny błąd losowy, którego źródłem jest losowy wybór próby. Z kolei obciążenie estymatora reprezentuje systematyczny błąd estymacji. Gdy zwiększamy liczebność próby, wtedy zmniejszeniu ulega właśnie $D(t_S)$, co oznacza , że zwiększa się precyzja oceny, ale obciążenie na ogół nie ulega zmianie. Wynika stąd, że oceny parametrów mogą być bardzo precyzyjne, a zarazem niedokładne. Szerzej problematyka oceny precyzji i dokładności w badaniach reprezentacyjnych potraktowana jest w mojej książce[8] na s. 31-36. Fakt ten należy wziąć pod uwagę już w czasie planowania badania reprezentacyjnego, a także analizy uzyskanych wyników.

Na s. 39 rozważany jest przedział ufności jako alternatywny sposób oceny precyzji parametru populacji. Autor słusznie pomija wywody prowadzące do budowy przedziału ufności i bezpośrednio stwierdza, że wyznaczenie przedziału polega na określeniu takich dwóch statystyk t_{1S} it_{2S} , że spełnione jest wyrażenie: $P(t_{1S} < \theta < t_{2S}) \geqslant \gamma$, gdzie przez γ oznaczono poziom ufności, czyli prawdopodobieństwo, z jakim przedział ufności $[t_{1S}, t_{2S}]$ obejmuje wartość szacowanego parametru θ. Autor następnie przyjmuje pewne założenia odnośnie estymatorów oraz liczebności próby, aby dojść do określenie przedziału ufności. W praktyce jednak obliczany jest średni błąd standardowy, tj. $s(\bar{x})$, a następnie mnożony jest przez odpowiedni kwantyl rozkładu normalnego, gdy liczebność próby jest dostatecznie liczna, aby utworzyć przedział ufności. Stwierdza się następnie, że taki przedział pokrywa prawdziwą wartość szacowanego parametru θ. Stwierdzenie to nie jest prawdziwe, gdy występują błędy nielosowe, takie jak błędy odpowiedzi, pomiaru lub braki odpowiedzi. Mamy wtedy do czynienia z precyzją, a przedział ufności pokrywa nie prawdziwą, ale obserwowaną wartość szacowanego parametru. Zilustrowano tę zależność na poniższym rysunku.

[8] J. Kordos (1988), Dokładność danych statystycznych, PWE, Warszawa.



Dalej w rozdziale 3 rozważana jest lokalizacja proporcjonalna prób w warstwach, warunkowa minimalizacja wariancji, warunkowa minimalizacja kosztów obserwacji zmiennej oraz warstwowanie populacji po jej wylosowaniu. Obszernie potraktowana została estymacja na podstawie próby grupowej, gdzie rozważano jednostopniową prostą próbę grupową, próbę dwustopniową oraz predykcję na podstawie grupowego modelu nadpopulacji.

Na końcach rozdziałów zamieszczono pytania i zadania, które mogą ułatwić czytelnikowi sprawdzenie przyswajanych wiadomości.

Warto zaznaczyć, że metoda reprezentacyjna korzysta nie tylko *z dorobku rachunku prawdopodobieństwa i statystyki matematycznej*, ale w poważnym stopniu wykorzystuje elementy *prakseologii i ogólnej metodologii badań*.

Wybór odpowiedniego rozwiązania w konkretnych warunkach opiera się zarówno na *przesłankach teoretycznych*, których dostarcza odpowiednio zbudowany dla rozwiązania danego zagadnienia model matematyczny, jak również *musi uwzględnić istniejące realnie warunki*. Jednostronne potraktowanie danego problemu badawczego nie może prowadzić do zadawalających wyników.

Ogólnie wiadomo, że próbka powinna dostarczyć *ocen parametrów z dostateczną precyzją.* Należy więc ustalić jakie parametry będziemy szacowali oraz jaki powinien być stopień ich precyzji (inaczej mówiąc dopuszczalne wielkości błędów losowych).

Teoretycznie badanie reprezentacyjne powinno być tak zaprojektowane, *aby dostarczyć ocen o minimalnych błędach (tj. maksymalnej precyzji), gdy ustalony jest ogólny koszt badania*; lub dla *ustalonej precyzji zminimalizować ogólny koszt badania.* Liczebność próbki spełniająca te warunki nazywa się *optymalną liczebnością próbki.*

W praktyce powinny być również wzięte pod uwagę inne okoliczności, takie jak *istnienie błędów nielosowych i innych obciążeń danych*. Błędy nielosowe zwykle wzrastają, od pewnego momentu, wraz ze zrostem liczebności próbki. Powinno się więc w praktyce poszukiwać kompromisu, uwzględniającego zarówno wielkości błędów losowych jak i błędów nielosowych. Uzyskanie tego kompromisu nie jest łatwe i wymaga głębokich prac badawczych.

Każde badanie statystyczne wymaga pewnego nakładu profesjonalnej pracy potrzebnej do jego przygotowania, realizacji w terenie i opracowania wyników. Już na etapie planowania i przygotowania badania reprezentacyjnego niezbędny jest udział specjalisty z metody reprezentacyjnej, posiadającego doświadczenia w przygotowaniu i prowadzeniu badań reprezentacyjnych. Nie wystarczy jednak sama znajomość teorii

metody reprezentacyjnej, ale niezbędna jest praktyka w projektowaniu i realizacji badań reprezentacyjnych oraz stała współpraca ze specjalistami z innych dziedzin.

Specjalista z metody reprezentacyjnej (MR) jest odpowiedzialny głównie za problemy związane z projektowaniem próbki, które obejmują lokalizacje wielkości próbki na różnych stopniach losowania oraz wybór metod i procedur losowania, jednostek losowania, identyfikacją, warstwowaniem i lokalizacją w warstwach.

Klasa zagadnień, którą nazwano "*wspólnym planem*" dotyczy również losowania, lecz decyzje te muszą być podjęte wspólnie ze specjalistami z danej dziedziny badań (ekonomistami, socjologami, agronomami, biologami itd. - w skrócie DD). Za określenie elementów populacji i zakresu badania powinien być odpowiedzialny specjalista DD, lecz wiedza i doradztwo MR powinna być tu wykorzystana, aby uzyskać dobry schemat losowania .

Proces estymacji jest często łączony z wyborem próbki, jako wspólne aspekty planu próbki (a nieobciążone estymatory, na przykład, mogą być tylko określone przez łączne rozważania prawdopodobieństw wyboru z wagami użytymi w estymatorze).

Błędy losowe zależą również w istotny sposób od schematu wyboru próby i teoria losowania jest tu niezbędna (wkład MR) lecz wkład DD - odpowiedzialnych za prezentację i analizę wyników - jest tu również niezbędny dla właściwej prezentacji błędów losowych. Dostępne środki finansowe i żądana precyzja wyników z próbki zależą od DD, lecz MR może również mieć na nie wpływ, ponieważ konflikty zwykle rozwijają się między żądanymi celami i ograniczonymi środkami.

Metoda reprezentacyjna, w sensie losowego wyboru próby umożliwiającej uogólnienie uzyskanych wyników z próby na całą badaną populację ze znacznymi oporami zdobywała sobie uznanie[9]. Rozwijała się nie tylko teoria badań reprezentacyjnych, ale także sztuka ich projektowania i realizacji w różnych dziedzinach badań w rozmaitych warunkach Wydaje się, że we wprowadzeniu do metody reprezentacyjnej należałoby uwzględnić nie tylko teorię metody reprezentacyjnej, ale także rzeczywiste zastosowania tej metody w praktyce Problemy te omawia szeroko S. Lohr w czasie rozważania zakresu szkolenia z metody reprezentacyjnej zarówno studentów jak i specjalistów z tego zakresu[10]. Warto w większym stopniu uwzględnić cytowany już podręcznik S. Lohr, a szczególnie jego drugie wydanie[11] z 2010 r., który uważam za najlepiej uwzględniający zarówno rozwój teorii, jak i praktyki metody reprezentacyjnej, od czasu opublikowania w 1953 r. podręcznika Hansena, Hurwitza i Madowa [12] Metoda reprezentacyjna mogłaby znaleźć szerokie zastosowanie właśnie w przygotowywanym w Polsce spisie ludności w 2011 r., w którym będą stosowane różne metody zbierania

[9] Kish, L. (1995), The Hundred Years' Wars of Survey Sampling, *Statistics in Transition,,* vol. 2, No 5, s. 813-830.

[10] Lohr, S. L. (2010), The 2009 Morris Hansen Lecture: The Care, Feeding, and Training of Survey Statisticians, *Journal of Official Statistics,* Vol. 26, No. 3, s. 395-409.

[11] Lohr, S.L. (2010), *Sampling: Design and Analysis*, Brooks/Cole, Cengage Learning.

[12] Hansen, M.H., Hurwitz, W.N., and Madow, W.G. (1953), *Sample Survey Methods and Theory,* Vol. 1 and 2, New York:Wiley.

danych z wykorzystaniem rejestrów oraz badań częściowych. Przeprowadzone właściwie pokontrolne badania spisowe, zgodnie z zasadami zalecanymi przez UN Statistics Division[13], w którym w szerokim zakresie proponuje się wykorzystanie metody reprezentacyjnej, może zapewnić odpowiednią jakość uzyskanych wyników.

Ogólnie oceniam recenzowaną pracę dość wysoko, a moje uwagi dotyczą niektórych aspektów, które warto rozważyć, gdyby praca została ponownie wydana. Uważam, że już we wprowadzeniu do metody reprezentacyjnej należy uwzględnić nie tylko teorię związaną z wyborem próby i estymacją uzyskanych wyników, ale także inne aspekty metody reprezentacyjnej, takie jak planowanie badania, badanie jakości danych, niepodjęcie badań, imputacje, kalibrację i metody badania małych obszarów, z którymi mamy do czynienia w praktyce badań reprezentacyjnych.

Jan Kordos
Szkoła Główna Handlowa w Warszawie

[13] United Nations, Statistics Division (2010), *Post Enumeration Surveys*, Operational guidelines, Technical Report, New York, April 2010

PODZIĘKOWANIE RECENZENTOM

Redakcja pragnie serdecznie podziękować Recenzentom za podjęty trud opiniowania opracowań naukowych nadesłanych w 20010 r. do „Przeglądu Statystycznego”. W wielu przypadkach opinie Recenzentów pozwoliły na znaczne udoskonalenie opracowań przed ich publikacją. Dziękujemy za wnikliwość, staranność i terminowość recenzji.

Podziękowanie to składamy następującym osobom:

Prof. dr hab. Sławomir Dorosiewicz,
Prof. dr hab. Krzysztof Jajuga,
Prof. dr hab. Stanisław Maciej Kot,
Prof. dr hab. Władysław Milo,
Prof. dr hab. Magdalena Osińska,
Prof. dr hab. Emil Panek,
Prof. dr hab. Mirosław Szreder,
Prof. dr hab. Marek Walesiak,
dr hab. Andrzej Bąk,
dr hab. Tadeusz Bołt,
dr hab. Marek Gruszczyński,
dr hab. Kazimierz Krauze,
dr hab. Jerzy Marzec,
dr hab. Paweł Miłobędzki,
dr hab. Jerzy Ossowski,
dr hab. Krystyna Pruska,
dr hab. Jerzy Rembeza,
dr hab. Krystyna Strzała,
dr Beata Jackowska,
dr Grzegorz Krzykowski,
dr Kamila Migdał Najman,
dr Krzysztof Najman,
dr Anna Pajor,
dr Anna Zamojska.

KONGRES STATYSTYKI POLSKIEJ
Z OKAZJI JUBILEUSZU 100-LECIA
POLSKIEGO TOWARZYSTWA STATYSTYCZNEGO
Poznań
18-20 kwietnia 2012 roku

W roku 2012 przypada jubileusz 100-lecia Polskiego Towarzystwa Statystycznego, organizacji skupiającej przedstawicieli służb statystyki publicznej i środowisk akademickich, samorządu terytorialnego i gospodarczego, jednostek administracji rządowej, zainteresowanych teorią i praktyką badań statystycznych. PTS rozwija działalność naukową w dziedzinie teorii, metodologii i praktyki badań statystycznych oraz podejmuje wszelkie wysiłki w celu upowszechniania wiedzy statystycznej w społeczeństwie. Aktywnie współpracuje z towarzystwami statystycznymi w innych państwach oraz takimi organizacjami, jak Międzynarodowy Instytut Statystyczny, Bernoulli-Society for Mathematical Statistics and Probability, International Society for Quality of Life Research, International Society for Quality-of-Life Studies czy Międzynarodowa Federacja Towarzystw Klasyfikacyjnych (IFCS).

Doskonałą okazją dla uczczenia 100-letniej historii i bogatej tradycji Polskiego Towarzystwa Statystycznego może stać się Kongres Statystyki Polskiej organizowany w Poznaniu w dniach od 18 do 20 kwietnia 2012 r.

Przewodnicząca Komitetu Organizacyjnego
dr hab. Elżbieta Gołata prof. nadzw. UEP: e-mail: elzbieta.golata@ue.poznan.pl

Sekretariat Komitetu Organizacyjnego
Urząd Statystyczny w Poznaniu, ul. J. H. Dąbrowskiego 79, 60-959 Poznań: e-mail: kongres2012@stat.gov.pl
tel. 61 27 98 266, 61 27 98 325 (pon-pt: 8-15), 61 27 98 343 (pon: 9-12, wt: 7-10, śr: 7-10), fax. +48 61 27 98 101

Ramowy program tego Kongresu obejmuje szereg sesji tematycznych, w tym sesję jubileuszową (historyczną), a także sesje poświęcone: metodologii badań statystycznych, statystyce regionalnej, statystyce ludności, statystyce społecznej i gospodarczej, problematyce danych statystycznych, statystyce zdrowia, sportu i turystyki. Zorganizowane zostaną również dwa panele dyskusyjne, koncentrujące się na tematach:

– Podstawowe problemy statystyki we współczesnym świecie;

– Przyszłość statystyki.

Organizatorami Kongresu są Polskie Towarzystwo Statystyczne, Główny Urząd Statystyczny, Urząd Statystyczny w Poznaniu oraz Uniwersytet Ekonomiczny w Poznaniu.

Zapraszamy do aktywnego włączenia się w obchody tego wyjątkowego święta Statystyki Polskiej i udziału w Kongresie Statystyki Polskiej. Aktualne informacje o kongresie zamieszczane są na stronie internetowej: http://www.stat.gov.pl/pts/kongres2012/ , gdzie dostępny jest również formularz zgłoszeniowy.

Przewodnicząca Komitetu Organizacyjnego:
dr hab. Elżbieta Gołata prof. nadzw. UEP: e-mail: elzbieta.golata@ue.poznan.pl

Sekretariat Komitetu Organizacyjnego
Urząd Statystyczny w Poznaniu
ul. J. H. Dąbrowskiego 79, 60-959 Poznań: e-mail: kongres2012@stat.gov.pl
tel. +48 61 27 98 325 - (pon-pt: 8-15),
+48 61 27 98 343 - (pon: 9-12, wt: 7-10, śr: 7-10)
+48 61 27 98 266
fax. +48 61 27 98 101

SPIS TREŚCI

SPRAWOZDANIA

RECENZJE