

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

RESEARCH PAPERS

of Wrocław University of Economics

Nr 426

Taksonomia 26

**Klasyfikacja i analiza danych –
teoria i zastosowania**



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2016

Redaktor Wydawnictwa: Agnieszka Flasińska

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Beata Mazur

Projekt okładki: Beata Dębska

Tytuł dofinansowany ze środków Narodowego Banku Polskiego
oraz ze środków Sekcji Klasyfikacji i Analizy Danych PTS

Informacje o naborze artykułów i zasadach recenzowania
znajdują się na stronach internetowych
www.pracnaukowe.ue.wroc.pl
www.wydawnictwo.ue.wroc.pl

Publikacja udostępniona na licencji Creative Commons
Uznanie autorstwa-Użycie niekomercyjne-Bez utworów zależnych 3.0 Polska
(CC BY-NC-ND 3.0 PL)



© Copyright by Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2016

ISSN 1899-3192 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)
e-ISSN 2392-0041
ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana

Zamówienia na opublikowane prace należy składać na adres:
Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
ul. Komandorska 118/120, 53-345 Wrocław
tel./fax 71 36 80 602; e-mail:econbook@ue.wroc.pl
www.ksiegarnia.ue.wroc.pl

Druk i oprawa: TOTEM

Spis treści

Wstęp	9
Jacek Batóg: Identyfikacja obserwacji odstających w analizie skupień / Influence of outliers on results of cluster analysis	13
Andrzej Bąk: Porządkowanie liniowe obiektów metodą Hellwiga i TOPSIS – analiza porównawcza / Linear ordering of objects using Hellwig and TOPSIS methods – a comparative analysis.....	22
Grażyna Dehnel: <i>MM</i> -estymacja w badaniu średnich przedsiębiorstw w Polsce / <i>MM</i> -estimation in the medium-sized enterprises survey in Poland.....	32
Andrzej Dudek: <i>Social network analysis</i> jako gałąź wielowymiarowej analizy statystycznej / Social network analysis as a branch of multidimensional statistical analysis.....	42
Iwona Foryś: Analiza dyskryminacyjna w wyborze obiektów podobnych w procesie szacowania nieruchomości / The discriminant analysis in selection of similar objects in the real estate valuation process	51
Gregory Kersten, Ewa Roszkowska, Tomasz Wachowicz: Ocena zgodności porządkowej systemu oceny ofert negocjatora z informacją preferencyjną / Analyzing the ordinal concordance of preferential information and resulting scoring system in negotiations.....	60
Iwona Konarzewska: Rankingi wielokryteriowe a współzależność liniowa kryteriów / Multi-criteria rankings and linear relationships among criteria	69
Anna Król, Marta Targaszewska: Zastosowanie klasyfikacji do wyodrębniania homogenicznych grup dóbr w modelowaniu hedonicznym / The application of classification in distinguishing homogeneous groups of goods for hedonic modelling.....	80
Marek Lubicz: Problemy doboru zmiennych objaśniających w klasyfikacji danych medycznych / Feature selection and its impact on classifier effectiveness – case study for medical data.....	89
Aleksandra Łuczak: Wpływ różnych sposobów agregacji opinii ekspertów w FAHP na oceny priorytetowych czynników rozwoju / Influence of different methods of the expert judgments aggregation on assessment of priorities for evaluation of development factors in FAHP.....	99
Iwona Markowicz: Tablice trwania firm w województwie zachodniopomorskim według rodzaju działalności / Companies duration tables in Zachodniopomorskie voivodship by the type of activity	108

Małgorzata Markowska, Danuta Strahl: Filary inteligentnego rozwoju a wrażliwość unijnych regionów szczebla NUTS 2 na kryzys ekonomiczny – analiza wielowymiarowa / Smart development pillars and NUTS 2 European regions vulnerability to economic crisis – a multidimensional analysis.....	118
Kamila Migdał-Najman, Krzysztof Najman: Hierarchiczne deglomeracyjne sieci SOM w analizie skupień / The hierarchical divisive SOM in the cluster analysis	130
Kamila Migdał-Najman, Krzysztof Najman: Hierarchiczne aglomeracyjne sieci SOM w analizie skupień / The hierarchical agglomerative SOM in the cluster analysis	139
Barbara Pawelek, Józef Pocięcha, Jadwiga Kostrzewska, Mateusz Baryła, Artur Lipieta: Problem wartości odstających w prognozowaniu zagrożenia upadłością przedsiębiorstw (na przykładzie przetwórstwa przemysłowego w Polsce) / Problem of outliers in corporate bankruptcy prediction (case of manufacturing companies in Poland)	148
Wojciech Roszka: Syntetyczne źródła danych w analizie przestrzennego zróżnicowania ubóstwa / Synthetic data sources in spatial poverty analysis.....	157
Małgorzata Rószkiewicz: Czynniki różnicujące efektywność pracy ankietera w wywiadach <i>face-to-face</i> w środowisku polskich gospodarstw domowych / Factors affecting the efficiency of face-to-face interviews with Polish households.....	166
Adam Sagan, Marcin Pelka: Analiza wielopoziomowa z wykorzystaniem danych symbolicznych / Multilevel analysis with application of symbolic data	174
Marcin Salamaga: Zastosowanie drzew dyskryminacyjnych w identyfikacji czynników wspomagających wybór kraju alokacji bezpośrednich inwestycji zagranicznych na przykładzie polskich firm / The use of classification trees in the identification of factors supporting the choice of FDI destination on the example of Polish companies.....	185
Agnieszka Stanimir: Pomiar wykluczenia cyfrowego – zagrożenia dla Pokolenia Y / Measurement of the digital divide – risks for Generation Y ...	194
Mirosława Sztemberg-Lewandowska: Grupowanie danych funkcjonalnych w analizie poziomu wiedzy maturzystów / Functional data clustering methods in the analysis of high school graduates' knowledge	206
Tadeusz Trzaskalik: Modelowanie preferencji w wielokryterialnych dyskretnych problemach decyzyjnych – przegląd bibliografii / Preference modeling in multi-criteria discrete decision making problems – review of literature	214

Joanna Trzęsiok: Metody nieparametryczne w badaniu zaufania do instytucji finansowych / Nonparametric methods in the study of confidence in financial institutions	226
Hanna Wdowicka: Analiza sytuacji na lokalnych rynkach pracy w Polsce / Local labour market analysis in Poland.....	235
Artur Zaborski: Zastosowanie skalowania dynamicznego oraz metody wektorów dryfu do badania zmian w preferencjach / The use of dynamic scaling and the drift vector method for studying changes in the preferences.....	245

Wstęp

W dniach 14–16 września 2015 r. w Hotelu Novotel Gdańsk Marina w Gdańsku odbyła się XXIV Konferencja Naukowa Sekcji Klasyfikacji i Analizy Danych PTS (XXIX Konferencja Taksonomiczna) „Klasyfikacja i analiza danych – teoria i zastosowania”, zorganizowana przez Sekcję Klasyfikacji i Analizy Danych Polskiego Towarzystwa Statystycznego oraz Katedrę Statystyki Wydziału Zarządzania Uniwersytetu Gdańskiego. Przewodniczącymi Komitetu Organizacyjnego konferencji byli prof. dr hab. Mirosław Szreder oraz dr hab. Krzysztof Najman, prof. nadzw. UG, sekretarzami naukowymi dr hab. Kamila Migdał-Najman, prof. nadzw. UG oraz dr hab. Anna Zamojska, prof. nadzw. UG, a sekretarzem organizacyjnym Anna Nowicka z Fundacji Rozwoju Uniwersytetu Gdańskiego.

Konferencja Naukowa została dofinansowana ze środków Narodowego Banku Polskiego.

Zakres tematyczny konferencji obejmował takie zagadnienia, jak:

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).

Zasadniczymi celami konferencji SKAD były 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 81 osób. Byli to pracownicy oraz doktoranci następujących uczelni i instytucji: AGH w Krakowie, Politechniki Łódzkiej, Politechniki Gdańskiej, Politechniki Opolskiej, Politechniki Wrocławskiej, Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie, Szkoły Głównej Handlowej w Warszawie, Uniwersytetu im. Adama Mickiewicza w Poznaniu, Uniwersytetu Ekonomicznego w Katowicach, Uniwersytetu Ekonomicznego w Krakowie, Uniwersytetu Ekonomicznego w Poznaniu, Uniwersytetu Ekonomicznego we Wrocławiu, Uniwersytetu Gdańskiego, Uniwersytetu Jana Kochanowskiego w Kielcach, Uniwersytetu Łódzkiego, Uniwersytetu Mikołaja Kopernika w Toruniu, Uniwersytetu Przyrodniczego w Poznaniu, Uniwersytetu Szczecińskiego, Uniwer-

sytetu w Białymstoku, Wyższej Szkoły Bankowej w Toruniu, a także przedstawiciele NBP i PBS Sp. z o.o.

W trakcie dwóch sesji plenarnych oraz trzynastu sesji równoległych wygłoszono 58 referatów poświęconych aspektom teoretycznym i aplikacyjnym zagadnienia klasyfikacji i analizy danych. Odbyła się również sesja plakatowa, na której zaprezentowano 14 plakatów. Obradom w poszczególnych sesjach konferencji przewodniczyli profesorowie: Józef Pocięcha, Eugeniusz Gatnar, Tadeusz Trzaskalik, Krzysztof Jajuga, Marek Walesiak, Barbara Pawełek, Feliks Wysocki, Ewa Roszkowska, Andrzej Sokołowski, Andrzej Bąk, Tadeusz Kufel, Mirosław Krzyśko, Krzysztof Najman, Małgorzata Rószkiewicz, Mirosław Szreder.

Teksty 25 recenzowanych artykułów naukowych stanowią zawartość prezentowanej publikacji z serii „Taksonomia” nr 26. Pozostałe recenzowane artykuły znajdują się w „Taksonomii” nr 27.

W pierwszym dniu konferencji odbyło się posiedzenie członków Sekcji Klasyfikacji i Analizy Danych Polskiego Towarzystwa Statystycznego, któremu przewodniczył prof. dr hab. Józef Pocięcha. Ustalono plan przebiegu zebrania obejmujący następujące punkty:

- A. Sprawozdanie z działalności Sekcji Klasyfikacji i Analizy Danych PTS.
- B. Informacje dotyczące planowanych konferencji krajowych i zagranicznych.
- C. Organizacja konferencji SKAD PTS w latach 2016 i 2017.
- D. Wybór przedstawiciela Rady Sekcji SKAD PTS do IFCS.
- E. Dyskusja nad kierunkami rozwoju działalności Sekcji.

Prof. dr hab. Józef Pocięcha otworzył posiedzenie Sekcji SKAD PTS. Sprawozdanie z działalności Sekcji Klasyfikacji i Analizy Danych PTS przedstawiła sekretarz naukowy Sekcji dr hab. Barbara Pawełek, prof. nadzw. UEK. Poinformowała, że obecnie Sekcja liczy 231 członków. Przypomniała, że na stronie internetowej Sekcji znajdują się regulamin, a także deklaracja członkowska. Poinformowała, że zostały opublikowane zeszyty z serii „Taksonomia” nr 24 i 25 (PN UE we Wrocławiu nr 384 i 385). W „Przeglądzie Statystycznym” (zeszyt 4/2014) ukazało się sprawozdanie z ubiegłorocznej konferencji SKAD, która odbyła się w Międzyzdrojach, w dniach 8–10 września 2014 r. Prof. Barbara Pawełek przedstawiła także informacje dotyczące działalności międzynarodowej oraz udziału w ważnych konferencjach członków i sympatyków SKAD.

W konferencji Międzynarodowego Stowarzyszenia Towarzystw Klasyfikacyjnych (IFCS – International Federation of Classification Societies) w dniach 6–8 lipca 2015 r. w Bolonii, zorganizowanej przez Università di Bologna, udział wzięło 19 osób z Polski (w tym 17 członków Sekcji), które wygłosiły 15 referatów (wkład członków SKAD – 79,0%). Ponadto prof. Józef Pocięcha był członkiem Komitetu Naukowego Konferencji z ramienia SKAD, członkiem Międzynarodowego Komitetu Nagród IFCS oraz organizatorem i przewodniczącym sesji nt. „Classification models for forecasting of economic processes”.

W konferencji „European Conference on Data Analysis” (Colchester, 2–4 września 2015 r.) zorganizowanej przez The German Classification Society (GfKI) we współpracy z The British Classification Society (BCS) i Sekcją Klasyfikacji i Analizy Danych PTS (SKAD) udział wzięło 18 osób z Polski (w tym 14 członków Sekcji), które wygłosiły 15 referatów (wkład członków SKAD – 66,0%). Ponadto profesorowie Krzysztof Jajuga oraz Józef Pociecha byli członkami Komitetu Naukowego konferencji, prof. Andrzej Dudek został poproszony przez organizatorów o przygotowanie referatu i wygłoszenie na Sesji Plenarnej „Cluster analysis in XXI century, new methods and tendencies”, prof. Krzysztof Jajuga był przewodniczącym sesji plenarnej, przewodniczącym sesji nt. „Finance and economics II” oraz organizatorem i przewodniczącym sesji nt. „Data analysis in finance”, prof. Józef Pociecha był organizatorem i przewodniczącym sesji nt. „Outliers in classification procedures – theory and practice”, prof. Andrzej Dudek był przewodniczącym sesji nt. „Machine learning and knowledge discovery II”.

Kolejny punkt posiedzenia Sekcji obejmował zapowiedzi najbliższych konferencji krajowych i zagranicznych, których tematyka jest zgodna z profilem Sekcji. Prof. dr hab. Józef Pociecha poinformował o dwóch wybranych konferencjach krajowych (były to XXXIV Konferencja Naukowa „Multivariate Statistical Analysis MSA 2015”, Łódź, 16–18 listopada 2015 r. i X Międzynarodowa Konferencja Naukowa im. Profesora Aleksandra Zeliasia nt. „Modelowanie i prognozowanie zjawisk społeczno-gospodarczych”, Zakopane, 10–13 maja 2016 r.) oraz o trzech wybranych konferencjach zagranicznych. Konferencja „European Conference on Data Analysis” odbędzie się na Uniwersytecie Ekonomicznym we Wrocławiu w dniach 26–28 września 2017 r. W przeddzień tej konferencji, tj. 25.09.2017 r., odbędzie się Niemiecko-Polskie Sympozjum nt. „Analizy danych i jej zastosowań GPSDAA 2017”. Następną konferencją Międzynarodowego Stowarzyszenia Towarzystw Klasyfikacyjnych (IFCS) odbędzie się w 2017 r. w Tokio. W 2019 r. Niemiecko-Polskie Sympozjum nt. „Analizy danych i jej zastosowań GPSDAA 2019” organizuje prof. Andreas Geyer-Schultz w Karlsruhe.

W następnym punkcie posiedzenia podjęto kwestię organizacji kolejnych konferencji SKAD. SKAD 2016 zorganizuje Katedra Metod Statystycznych Wydziału Ekonomiczno-Socjologicznego Uniwersytetu Łódzkiego.

W kolejnej części zebrania dokonano wyboru przedstawiciela Rady Sekcji SKAD PTS do IFCS na kadencję 2016–2019. Powołano Komisję Skrutacyjną, której przewodniczącym został prof. Tadeusz Kufel, a członkami dr hab. Iwona Konarzewska i dr Dominik Rozkrut. Profesor Józef Pociecha poprosił zebranych o proponowanie kandydatur zgłaszając jednocześnie prof. Andrzeja Sokołowskiego. Wobec braku następnych kandydatur listę zamknięto. Komisja Skrutacyjna przeprowadziła głosowanie tajne. W głosowaniu uczestniczyło 41 członków Sekcji. Profesor Andrzej Sokołowski został przedstawicielem Rady Sekcji SKAD PTS do

IFCS na kadencję 2016–2019, uzyskując następujący wynik: 39 głosów na „tak”, 1 głos na „nie”, 1 głos był nieważny.

W ostatnim punkcie zebrania dyskutowano nad kierunkami rozwoju działalności Sekcji obejmującymi następujące problemy: udział w międzynarodowym ruchu naukowym (wspólne granty, publikacje), umiędzynarodowienie konferencji SKAD (uczestnicy zagraniczni, dwujęzyczność konferencji), wydawanie własnego czasopisma.

Profesor Józef Pociecha zamknął posiedzenie Sekcji SKAD.

Krzysztof Jajuga, Marek Walesiak

Iwona Markowicz

Uniwersytet Szczeciński
e-mail: iwona.markowicz@wneiz.pl

**TABLICE TRWANIA FIRM
W WOJEWÓDZTWIE ZACHODNIOPOMORSKIM
WEDŁUG RODZAJU DZIAŁALNOŚCI**

**COMPANIES DURATION TABLES
IN ZACHODNIOPOMORSKIE VOIVODSHIP
BY THE TYPE OF ACTIVITY**

DOI: 10.15611/pn.2016.426.11

Streszczenie: Celem przeprowadzonych badań było skonstruowanie tablic trwania firm i analiza przebiegu funkcji intensywności ich likwidacji. Jest to model nieparametryczny należący do metod analizy przeżycia (trwania) zjawisk. Podmioty gospodarcze, jako ważny element gospodarki kraju, są przedmiotem coraz liczniejszych analiz prezentowanych w literaturze polskiej i zagranicznej. W niniejszym badaniu wykorzystano dane z rejestru REGON, dotyczące firm powstałych w województwie zachodniopomorskim w latach 2009–2011. Podmioty te obserwowane były do końca 2013 roku. Nowym przeprowadzonych badań jest analiza według rodzaju działalności. Tablice trwania zbudowano dla 13 sekcji PKD2007. Ogółem analizie trwania poddano 59 587 powstałych firm. Spośród nich do końca 2013 r. zlikwidowanych zostało 22 234 podmioty. Pozostałe firmy przetrwały poza okres obserwacji i zostały sklasyfikowane jako cenzurowane i tak ujęte w tablicach trwania. W tablicach tych zawarto funkcję hazardu (funkcję intensywności likwidacji firm), która, jak podkreśla się w literaturze anglojęzycznej, zazwyczaj ma kształt odwróconej litery U (co jest zgodne z teoretycznym modelem uczenia się).

Słowa kluczowe: tablice trwania firm, analiza trwania, funkcja hazardu (funkcja intensywności likwidacji firm).

Summary: The aim of the study was to construct the tables of companies duration. This is a non-parametric model belonging to the method of the duration analysis. Economic entities, as an important part of the country economy, are the subject of a growing number of analyses presented in literature. This study used data from the registry of REGON, related to companies established in Zachodniopomorskie voivodship in 2009-2011. These entities were observed to the end of 2013. Novelty of these researches is the analysis by type of

activity. Duration tables were constructed for 13 section of PKD2007. Generally, 59 587 of established companies were analyzed. 22 234 of these companies were liquidated to the end of 2013. Other firms have survived beyond the observation period and have been classified as censored. The tables include hazard function (intensity function of the companies liquidation) which, as highlighted in the English literature, usually have a shape of inverted U (which is compatible with the theoretical model of learning).

Keywords: Duration tables, duration analysis, hazard function.

1. Wstęp

Podmioty gospodarcze (firmy) są przedmiotem coraz liczniejszych analiz prezentowanych w literaturze polskiej i zagranicznej. Analiza żywotności firm jest połączeniem metod statystycznych analizy wskaźnikowej i analizy trwania w celu wszechstronnego badania stanu, struktury, dynamiki, natężenia i modelowania czasu trwania zbiorowości firm [Markowicz 2012]. Zbiorowości firm, podobnie jak kohorty ludzi, można badać wykorzystując metody analizy przeżycia (trwania).

Problematykę badawczą dotyczącą trwania firm w polskiej literaturze naukowej można podzielić na: analizę wskaźników i modelowanie trwania firm [Bławat (red.) 2004; Dehnel 2010; Dominiak 2005; Gołata 2009; Hozer, Markowicz 2002; Hozer, Markowicz, Stolorz 2008; Markowicz 2000, 2012; Poznńska 2007] oraz analizę i predykcję bankructwa firm [Appenzeller 2004; Hadasik 1998; Mączyńska (red.) 2008; Nehrebecka, Dzik 2013; Pocięcha 2010; Pocięcha, Pawełek 2011; Wędzki 2005]. Podział ten wynika ze sposobu zakończenia działalności przez przedsiębiorstwo oraz wykorzystywanie danych finansowych w analizie bankructwa.

Celem przeprowadzonych badań było skonstruowanie tablic trwania firm i analiza przebiegu funkcji intensywności likwidacji firm powstałych w woj. zachodniopomorskim w latach 2009–2011. Obserwacją objęto 13 sekcji PKD.

2. Tablice trwania firm i intensywność likwidacji

W analizie trwania wyróżnia się modele parametryczne, semiparametryczne i nieparametryczne. Budowa modeli parametrycznych wymaga przyjęcia teoretycznego rozkładu badanej zmiennej [Frątczak, Gach-Ciepiela, Babiker 2005; Gazińska 2003], co w przypadku badania firm nie jest możliwe. Stąd też powszechniej stosowane są modele nieparametryczne, których przykładem są tablice trwania. Jest to model tabelaryczny, najczęściej stosowany w demogra-

fii, jako tablice trwania życia (wymieralności). Wówczas na ogół przyjmowany jest roczny przedział wieku, a tablice takie nazywa się tablicami pełnymi. W badaniu firm przyjęto trzymiesięczny okres dla tablic kohortowych. Kohortą jest grupa osób lub obiektów, wyodrębniona na podstawie wspólnie przeżytego zdarzenia w ściśle określonym miejscu i czasie. Tablice kohortowe są wynikiem obserwacji kohorty od momentu jej powstania, przez długi okres, do momentu opuszczenia kohorty przez wszystkie obiekty lub do zakończenia obserwacji, co się wiąże z uwzględnieniem jednostek cenzurowanych.

Tablice kohortowe firm prezentują proces zmian zachodzących w składzie kohorty z upływem czasu. Kohortę tworzą firmy powstałe w danym okresie. Rejestracji podlegają fakt likwidacji firmy, a więc wyjście z kohorty, oraz czas tego zdarzenia, licząc od początku obserwacji (rejestracji). Tablice kohortowe należą do modeli o czasie ciągłym, ale przedstawione oszacowania funkcji są dyskretne, dlatego wartość zmiennej czasowej, pogrupowanej w jednakowe przedziały, podano jako początek przedziału (t). Liczbę firm dotrwałych n_t podano na początek okresu i wyznaczono jako $n_t = n_{t-1} - (z_{t-1} + c_{t-1})$, przy czym n_t dla $t = 0$ (n_0) oznacza początkową liczebność kohorty. Przez z_t oznaczono liczbę firm zlikwidowanych w przedziale $\langle t, t + 1 \rangle$, czyli liczbę jednostek, które doznały analizowanego zdarzenia w danym przedziale czasu, a przez c_t – liczbę firm, które nie doświadczyły zdarzenia do końca okresu obserwacji kohorty (są to firmy, dla których znany jest czas do końca obserwacji, a nie czas trwania; mają one jednak wpływ na wyznaczone w tablicy funkcje).

Kolejną wielkość w kohortowej tablicy trwania oszacowano w ujęciu dyskretnym, gdyż może być wyznaczona tylko dla przedziału czasu. Jest to prawdopodobieństwo likwidacji firmy w przedziale czasu f_t , definiowane jako warunkowe prawdopodobieństwo likwidacji firmy w przedziale czasu trwania $\langle t, t + 1 \rangle$, pod warunkiem że firma nie została zlikwidowana do czasu t . Rozkład trwania firm nie może być przyporządkowany żadnemu ze znanych typów rozkładu prawdopodobieństwa, dlatego funkcje opisujące proces trwania podmiotów gospodarczych nie są znane, a w tablicach trwania przedstawione są ich oszacowania wyznaczone na podstawie danych empirycznych. Estymatorem prawdopodobieństwa likwidacji firmy w przedziale czasu \hat{f}_t jest stosunek liczby firm zlikwidowanych w danym przedziale czasu (z_t) do liczby firm, które dotrwały do początku przedziału (n_t).

Prawdopodobieństwo przetrwania S_t oraz intensywność likwidacji firm h_t to funkcje z natury ciągłe, ale w tablicach są przedstawiane w ujęciu dyskretnym. Prawdopodobieństwo przetrwania S_t , wyznaczone dla przedziału $\langle t, t + 1 \rangle$, jest prawdopodobieństwem, tego że firma zostanie zlikwidowana po czasie $t + 1$. Estymatorem funkcji prawdopodobieństwa przetrwania \hat{S}_t jest stosunek liczby firm, które dotrwały do czasu $t + 1$, do wyjściowej liczebności kohorty.

Estymator \hat{h}_t funkcji intensywności jest wyznaczany jako stosunek estymatora prawdopodobieństwa likwidacji firmy w przedziale $\langle t, t + 1 \rangle$ do połowy sumy estymatorów prawdopodobieństw przetrwania dla przedziałów $\langle t, t + 1 \rangle$ i $\langle t - 1, t \rangle$

[Balicki 2006]: $\hat{h}_t = \frac{\hat{f}_t}{(\hat{S}_t + \hat{S}_{t-1})/2}$. Wiele badań potwierdza, że funkcja intensywności likwidowania przedsiębiorstw ma kształt odwróconej litery U z określonym maksimum, co jest zgodne z teoretycznym modelem uczenia się (więcej w: [Markowicz 2012]).

3. Dane statystyczne

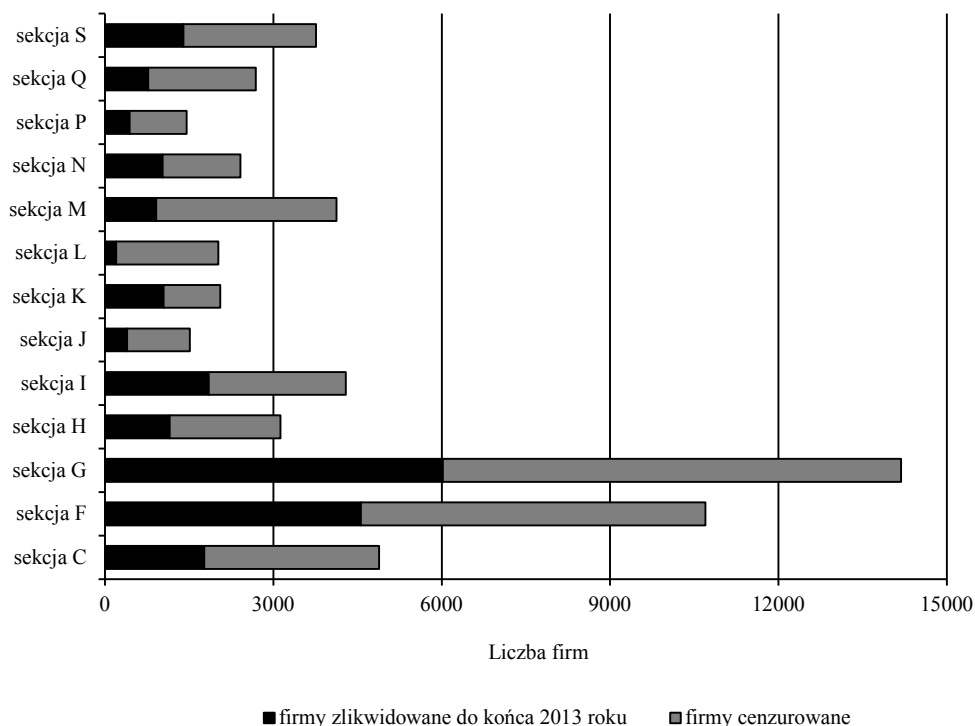
W artykule zaprezentowane zostały wyniki analizy kohortowej. Kohorty stanowią firmy danej sekcji (PKD2007) powstałe w województwie zachodniopomorskim w latach 2009–2011 i obserwowane do końca 2013 r. Informacje o liczbie

Tabela 1. Liczba firm powstałych w latach 2009–2011 i zlikwidowanych do końca 2013 r. w województwie zachodniopomorskim według sekcji PKD2007

Sekcje	Liczba firm	
	powstałych w latach 2009–2011	zlikwidowanych do końca 2013 r.
Ogółem 21 sekcji	59 587	22 234
Razem 13 sekcji badanych	57 209	21 531
Sekcja C, przetwórstwo przemysłowe	4 885	1 767
Sekcja F, budownictwo	10 696	4 557
Sekcja G, handel hurtowy i detaliczny; naprawa pojazdów samochodowych, włączając motocykle	14 182	6 018
Sekcja H, transport i gospodarka magazynowa	3 129	1 155
Sekcja I, działalność związana z zakwaterowaniem i usługami gastronomicznymi	4 292	1 848
Sekcja J, informacja i komunikacja	1 514	391
Sekcja K, działalność finansowa i ubezpieczeniowa	2 056	1 044
Sekcja L, działalność związana z obsługą rynku nieruchomości	2 020	198
Sekcja M, działalność profesjonalna, naukowa i techniczna	4 124	915
Sekcja N, działalność w zakresie usług administrowania i działalność wspierająca	2 412	1 028
Sekcja P, edukacja	1 452	442
Sekcja Q, opieka zdrowotna i pomoc społeczna	2 687	768
Sekcja S, pozostała działalność usługowa	3 760	1 400

Źródło: opracowanie własne na podstawie danych z rejestru REGON.

powstałych i zlikwidowanych firm według sekcji zawarto w tab. 1. Natomiast na rys. 1 przedstawiono liczbę firm zlikwidowanych i cenzurowanych (niezlikwidowanych do końca obserwacji) według sekcji. W latach 2009–2011 w województwie zachodniopomorskim powstało 59 587 firm. Spośród 21 sekcji analizie poddano 13. Pominięto sekcje o stosunkowo małej liczbie podmiotów. Zarówno wśród firm powstających jak i likwidowanych przeważały firmy handlowe i budowlane.



Rys. 1. Liczba firm powstałych w latach 2009–2011, zlikwidowanych i cenzurowanych w województwie zachodniopomorskim – obserwacja do końca 2013 roku

Źródło: opracowanie własne na podstawie danych z rejestru REGON.

4. Wyniki badań

Dla 59 587 firm powstałych w województwie zachodniopomorskim w latach 2009–2011 skonstruowano tablicę trwania (tab. 2). Czas trwania analizowany jest w przedziałach. W tabeli podano początek przedziału czasu w miesiącach (t). Ta-

bele takie zbudowano również dla 13 sekcji PKD2007, a graficznie zaprezentowano oceny intensywności likwidacji firm (funkcji hazardu) ogółem i według sekcji (rys. 2).

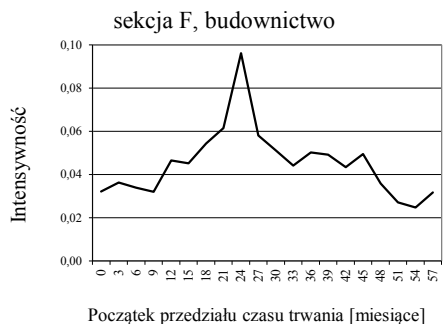
Tabela 2. Kohortowa tablica trwania firm powstałych w województwie zachodniopomorskim w latach 2009–2011 – obserwacja do końca 2013 roku (fragment)

Czas trwania [miesiące]	Liczba firm, które dotrwały	Liczba firm zlikwidowanych w przedziale czasu	Liczba firm cenzurowanych w przedziale czasu	Ocena prawdopodobieństwa likwidacji firmy w przedziale czasu	Ocena prawdopodobieństwa przetrwania	Ocena intensywności likwidacji firm
t	n_t	z_t	c_t	\hat{f}_t	\hat{S}_t	\hat{h}_t
					1,00000	
0	59 587	1592	0	0,02672	0,97328	0,02708
3	57 995	2128	0	0,03669	0,93757	0,03840
6	55 867	1617	0	0,02894	0,91043	0,03132
⋮	⋮	⋮	⋮	⋮	⋮	⋮
51	10 703	148	3716	0,01383	0,62861	0,02195
54	6 839	73	3842	0,01067	0,62739	0,01700
57	2 924	31	2893	0,01060	0,62686	0,01691

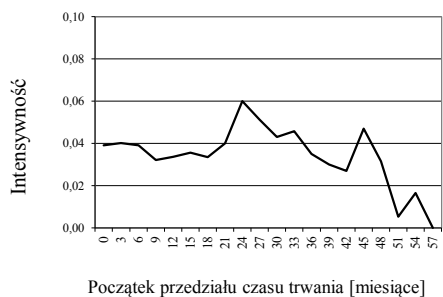
Źródło: obliczenia własne.

Analiza wykresów przedstawionych na rys. 2 wskazuje na przybliżony kształt odwróconej litery U z maksimum w przedziale od 24 do 27 miesięcy. Zatem po dwóch latach prowadzenia działalności gospodarczej intensywność likwidowania firm maleje. Sytuacja taka dotyczy wszystkich firm ogółem oraz większości sekcji. Nieco inny kształt przybiera funkcja intensywności dla sekcji H. Firmy transportowe osiągają próg po 2 i po 4 latach działalności, po czym intensywność maleje. Jest to zapewne związane z dużymi ponoszonymi kosztami początkowymi. Firmy zajmujące się zakwaterowaniem i gastronomią (sekcja I) swój próg osiągają szybko, bo już w przedziale od 3 do 6 miesięcy. Jest to związane z kończeniem działalności po sezonie letnim. Funkcje intensywności dla sekcji J (informacja i komunikacja) i L (działalność związana z obsługą rynku nieruchomości) najmniej przypominają omawiany kształt odwróconej litery U .

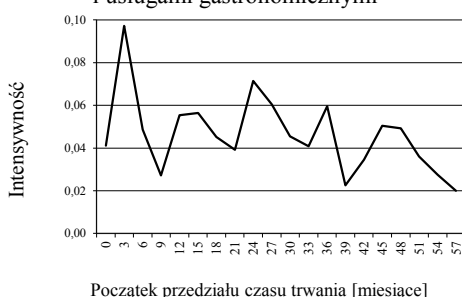
Przewidywany kolejny etap badań będzie dotyczył analizy przestrzennej, np. w powiatach województwa zachodniopomorskiego.



sekcja H, transport i gospodarka magazynowa



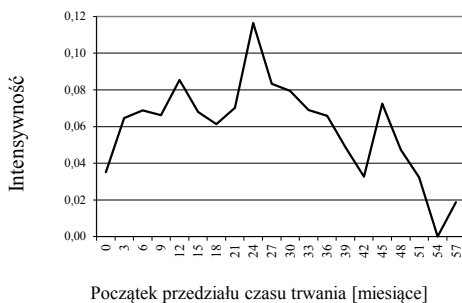
sekcja I, działalność związana z zakwaterowaniem i usługami gastronomicznymi

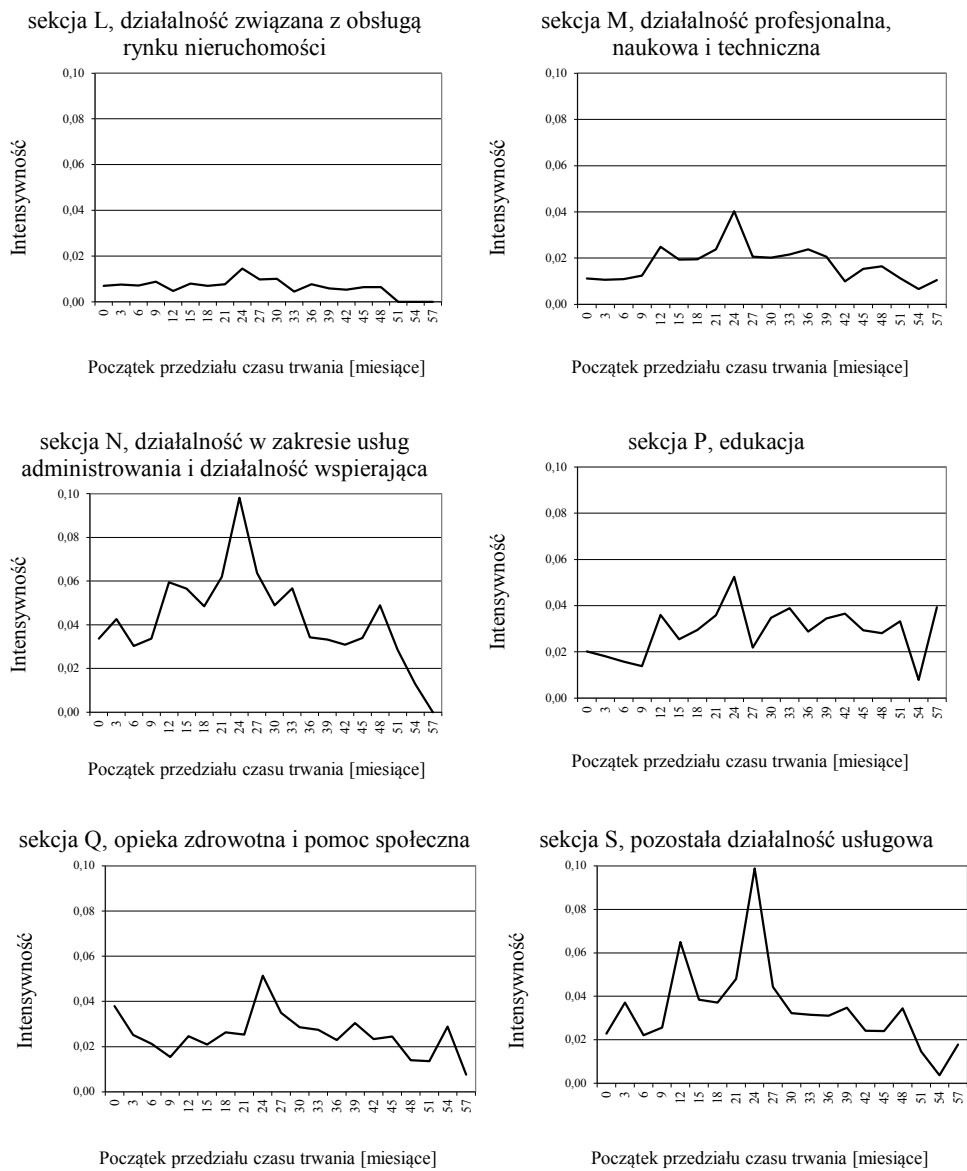


sekcja J, informacja i komunikacja



sekcja K, działalność finansowa i ubezpieczeniowa





Rys. 2. Ocena intensywności likwidacji firm według sekcji PKD2007

Źródło: opracowanie własne.

5. Zakończenie

Wyniki badań zaprezentowane w niniejszym artykule wskazują, że funkcje intensywności likwidacji podmiotów powstałych w województwie zachodniopomorskim w latach 2009–2011 według rodzaju działalności w większości przypominają typowy kształt odwróconej litery U. Maksimum funkcji intensywności, czyli pewien krytyczny moment jest osiągany po 24 miesiącach działalności. Wyjątki stanowią sekcje: H – transport i gospodarka magazynowa (spadek po 45 miesiącach), I – działalność związana z zakwaterowaniem i usługami gastronomicznymi (max w przedziale 3–6 miesięcy), J – informacja i komunikacja, L – działalność związana z obsługą rynku nieruchomości (intensywność wyrównana).

Literatura

- Appenzeller D., 2004, *Ekonometryczna analiza czynników kształtujących skalę i dynamikę upadłości w Polsce*, [w:] D. Appenzeller (red.), *Upadłości przedsiębiorstw w Polsce w latach 1990–2003. Teoria i praktyka*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań.
- Balicki A., 2006, *Analiza przeżycia i tablice wymieralności*, PWE, Warszawa.
- Bławat F. (red.), 2004, *Przetrwanie i rozwój przedsiębiorstw*, Scientific Publishing Group, Gdańsk.
- Dehnel G., 2010, *Rozwój mikroprzedsiębiorczości w Polsce w świetle estymacji dla małych domen*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań.
- Dominiak P., 2005, *Sektor MSP we współczesnej gospodarce*, Wydawnictwo Naukowe PWN, Warszawa.
- Frątczak E., Gach-Ciepiela U., Babiker H., 2005, *Analiza historii zdarzeń. Elementy teorii, wybrane przykłady zastosowań*, Oficyna Wydawnicza SGH, Warszawa.
- Gazińska M., 2003, *Potencjał demograficzny w regionie. Analiza ilościowa*, Rozprawy i Studia, t. 448, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin, s. 124–140.
- Gołata E., 2009, *Mikroprzedsiębiorstwa w badaniu SP3 oraz dodatkowych źródłach informacji*, [w:] J. Paradysz (red.), *Statystyka małych obszarów w badaniach podmiotów gospodarczych*, Zeszyty Naukowe Uniwersytetu Ekonomicznego w Poznaniu nr 116, s. 11–26.
- Hadasik D., 1998, *Upadłość przedsiębiorstw w Polsce i metody jej prognozowania*, Zeszyty Naukowe, Prace Habilitacyjne, z. 153, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań.
- Hozer J., Markowicz I., 2002, *Małe firmy. Analizy i diagnozy*, Rozprawy i Studia, t. 437, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.
- Hozer J., Markowicz I., Stolorz B., 2008, *Zastosowanie metod analizy historii zdarzeń w badaniu czasu funkcjonowania firm*, Zapol, Szczecin.
- Markowicz I., 2000, *Statystyczna analiza rentowności i żywotności małych firm w województwie szczecińskim*, praca doktorska, maszynopis, Szczecin.
- Markowicz I., 2012, *Statystyczna analiza żywotności firm*, Rozprawy i Studia, t. (CMIX) 835, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.
- Mączyńska E. (red.), 2008, *Bankructwa przedsiębiorstw. Wybrane aspekty instytucjonalne*, Oficyna Wydawnicza SGH, Warszawa.
- Nehrebecka N., Dzik A.M., 2013, *Zdolność przetrwania przedsiębiorstw w Polsce*, Wiadomości Statystyczne, nr 5.

- Pociecha J., 2010, *Metodologiczne problemy prognozowania bankructwa*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 107, Taksonomia 17, s. 54–61.
- Pociecha J., Pawełek B., 2011, *Prognozowanie bankructwa a koniunktura gospodarcza*, Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie, nr 873, Metody Analizy Danych, Kraków.
- Poznańska K., 2007, *Przetrwanie małych i średnich przedsiębiorstw – koncepcje i ich weryfikacja empiryczna w gospodarce polskiej*, [w:] T. Łuczka (red.), *Małe i średnie przedsiębiorstwa. Szkice o współczesnej przedsiębiorczości*, Wydawnictwo Politechniki Poznańskiej, Poznań.
- Wędzki D., 2005, *Zastosowanie logitowego modelu upadłości przedsiębiorstw*, *Ekonomista*, nr 5, s. 691–705.