

Artur Mikulec, Małgorzata Misztal

Uniwersytet Łódzki
e-mails: amikulec@uni.lodz.pl; mmisztal@uni.lodz.pl

ZASTOSOWANIE METODY REKURENCYJNEGO PODZIAŁU W ANALIZIE TRWANIA PRZEDSIĘBIORSTW WOJEWÓDZTWA ŁÓDZKIEGO

APPLICATION OF RECURSIVE PARTITIONING METHOD IN FIRMS' DURATION ANALYSIS IN LODZKIE VOIVODSHIP

DOI: 10.15611/pn.2018.507.18

JEL Classification: C10, C14, C41

Streszczenie: W artykule przedstawiono wyniki analizy czasu trwania przedsiębiorstw w województwie łódzkim na podstawie danych krajowego rejestru urzędowego podmiotów gospodarki narodowej REGON. W badaniach uwzględniono przedsiębiorstwa powstałe i zlikwidowane w województwie łódzkim w latach 2010-2015. Celem prowadzonych analiz była identyfikacja charakterystyk przedsiębiorstw, które w istotny sposób są związane z czasem trwania przedsiębiorstw oraz ryzykiem ich likwidacji. Do analizy czasu trwania wykorzystano metodę Kaplana-Meiera oraz model proporcjonalnego hazardu Coxa. Jako alternatywę dla tych metod zaproponowano połączenie metody rekurencyjnego podziału, w której dokonywany jest podział zbioru obiektów na jednorodne podzbiory, z estymacją funkcji przeżycia Kaplana-Meiera w każdym z węzłów końcowych. Graficzną prezentacją połączonych analiz są drzewa przeżycia. W pracy za pomocą dwóch algorytmów budujących drzewa przeżycia wyodrębniono i scharakteryzowano grupy przedsiębiorstw o podobnym czasie przeżycia.

Słowa kluczowe: przedsiębiorstwa, analiza trwania, metoda rekurencyjnego podziału, drzewa przeżycia.

Summary: The article presents the results of enterprises duration time analysis in Lodzkie Voivodship on the basis of data from the National Official Business Register REGON. The study examines enterprise births and deaths in Lodzkie Voivodship in the years 2010-2015. The goal of the analysis was to identify those characteristics of enterprises which were significantly related to the duration time and the risk of firm's liquidation. In the duration analysis, the Kaplan-Meier survival curves and Cox proportional-hazards regression model were used. As an alternative, recursive partitioning in combination with the Kaplan-Meier method is proposed. The model building process is based on recursive partitioning the set of enterprises into homogenous subsets and the Kaplan-Meier estimate of the survival function in each terminal node is reported. The results can be presented graphically using survival trees. In the paper

two tree-based algorithms were used to distinguish and describe groups of enterprises with similar duration time.

Keywords: enterprises, duration analysis, recursive partitioning method, survival trees.

1. Wstęp

Analiza trwania przedsiębiorstw obejmuje metody analizy danych, w których badaną zmienną jest czas do pojawienia się danego zdarzenia, np. czas od momentu powstania przedsiębiorstwa do jego likwidacji (por. np. [Balicki 2006]).

Typowym zagadnieniem w takich analizach jest poszukiwanie zmiennych istotnie wpływających na czas trwania przedsiębiorstwa. Występowanie obserwacji uciętych wyklucza zastosowanie tradycyjnych metod statystycznych, np. analizy regresji. Do najczęściej wykorzystywanych metod w analizie trwania należą: estymacja funkcji przeżycia metodą Kaplana-Meiera oraz model proporcjonalnego hazardu Coxa.

W polskiej literaturze przedmiotu znaleźć można kilka interesujących pozycji dotyczących modelowania czasu trwania przedsiębiorstw z wykorzystaniem obu wspomnianych metod. W pracach Markowicz i Stolorz [2006, 2007] analizowano krzywe przeżycia przedsiębiorstw powstałych w Szczecinie w latach 1990, 1994 i 1997, a w książce Hozera, Markowicz i Stolorz [2008] przedstawiono wyniki estymacji krzywej przeżycia Kaplana-Meiera oraz parametrów modelu regresji Coxa dla wyżej wymienionych przedsiębiorstw. Rozszerzone wyniki analizy trwania przedsiębiorstw powstałych w latach 2001-2006 w Szczecinie oraz w latach 2009–2011 w województwie zachodniopomorskim prezentowała w swoich pracach Markowicz [2012, 2015, 2017a, 2017b]. Na tle wspomnianych prac warto również wymienić rozprawę doktorską, w której analizie poddano mikroprzedsiębiorstwa w Polsce w latach 2002–2007 [Szymański 2011]. Trwaniem podmiotów gospodarczych (w tym przedsiębiorstw) w Małopolsce w latach 2002-2008 z zastosowaniem estymacji Kaplana-Meiera zajmował się Śmiech [2011], a swego rodzaju kontynuację tych badań stanowi analiza przedstawiona przez Gurgula i Zająca [2016]. Z kolei mikroprzedsiębiorstwa powstałe w Gdańsku w 2009 r. były przedmiotem analiz Jackowskiej [2015]. W pracach Mikulca [2017a, 2017b] znaleźć można rozszerzone wyniki analizy wskaźnikowej i analizy trwania przedsiębiorstw województwa łódzkiego w latach 2001–2015.

Jako alternatywę dla metody Kaplana-Meiera i modelu proporcjonalnego hazardu Coxa można zaproponować metodę rekurencyjnego podziału, której graficzną prezentacją są drzewa przeżycia (*survival trees*). W literaturze przedmiotu brak jest przykładów zastosowania metody rekurencyjnego podziału w analizie trwania przedsiębiorstw, stąd też autorzy podjęli w pracy próbę połączenia tych dwóch grup metod analizy.

W badaniu wykorzystano bazę danych przedsiębiorstw powstałych (w tym zlikwidowanych) w województwie łódzkim w latach 2010-2015. Celem prowadzonych analiz była identyfikacja charakterystyk przedsiębiorstw, które w istotny sposób są związane z czasem trwania przedsiębiorstw oraz ryzykiem ich likwidacji.

2. Charakterystyka danych i wykorzystanych metod analizy czasu trwania

W analizie trwania przedsiębiorstw¹ wykorzystano dane jednostkowe o 119,3 tys. przedsiębiorstwach powstałych w województwie łódzkim w latach 2010-2015. Spośród nich 28,3 tys. jednostek uległo likwidacji do końca 2015 r. (23,7%), a 91,0 tys., tj. 76,3%, funkcjonowało nadal (dane cenzurowane) – por. tab. 1.

Tabela 1. Przedsiębiorstwa w województwie łódzkim według roku powstania

Rok powstania	Przedsiębiorstwa powstałe		Przedsiębiorstwa zlikwidowane	Przedsiębiorstwa cenzurowane
	N	%		
OGÓLEM	119 297	100,0	28 321	90 976
2010	22 734	19,1	9 947	12 787
2011	19 270	16,1	7 219	12 051
2012	19 049	16,0	5 490	13 559
2013	19 940	16,7	3 729	16 211
2014	19 396	16,3	1 507	17 889
2015	18 908	15,8	429	18 479

Źródło: opracowanie własne.

¹ Przedsiębiorstwa zostały wybrane zgodnie z metodologią badania SP-3 *Sprawozdanie o działalności gospodarczej przedsiębiorstw* – rocznego badania przedsiębiorstw o liczbie pracujących do 9 osób, prowadzonego przez GUS i Urząd Statystyczny w Łodzi – przy świadomym pominięciu kryterium wielkości przedsiębiorstwa (liczby osób pracujących). Zamierzeniem autorów było, aby analiza dotyczyła ogółu przedsiębiorstw prowadzących działalność produkcyjną, handlową lub usługową na zasadach wolnorynkowych – w celach zarobkowych i na własny rachunek (bez względu na ich wielkość). Analizowaną zbiorowość stanowiły zatem jednostki zorganizowane w formie: spółek handlowych (osobowych i kapitałowych), spółek cywilnych, przedsiębiorstw państwowych, spółdzielni, oddziałów przedsiębiorców zagranicznych, osób fizycznych prowadzących działalność gospodarczą oraz państwowych jednostek organizacyjnych prowadzących działalność w sekcji: A (z wyłączeniem osób fizycznych prowadzących indywidualne gospodarstwa rolne), B, C, D, E, F, G, H, I, J (z wyłączeniem instytucji kultury mających osobowość prawną), K (z wyłączeniem banków, spółdzielczych kas oszczędnościowo-kredytowych, instytucji ubezpieczeniowych, biur i domów maklerskich, towarzystw i funduszy inwestycyjnych oraz towarzystw i funduszy emerytalnych), L, M, N, P (z wyłączeniem szkolnictwa wyższego), Q (z wyłączeniem samodzielnych publicznych zakładów opieki zdrowotnej), R (z wyłączeniem instytucji kultury mających osobowość prawną), S [*Program badań statystycznych...*, s. 377–379].

Dane jednostkowe o przedsiębiorstwach powstałych i zlikwidowanych w województwie łódzkim zostały pozyskane z Urzędu Statystycznego w Łodzi i zawierały: REGON², lokalizację, formę własności, formę prawną, PKD, odpowiednio datę powstania przedsiębiorstwa i/lub datę zakończenia działalności gospodarczej oraz wielkość przedsiębiorstwa – mierzona liczbą osób pracujących podaną przedziałowo (do 9, 10–49, 50 i więcej).

W przypadku formy własności, z uwagi na dużą liczbę jej wariantów, dokonano agregacji przedsiębiorstw w 3 podgrupy: jednostki sektora prywatnego (własność „czysta”), jednostki sektora publicznego (własność „czysta”), przedsiębiorstwa sektora prywatnego (własność mieszana). Zarówno wśród przedsiębiorstw zlikwidowanych, jak i cenzurowanych dominowały przedsiębiorstwa sektora prywatnego (własność „czysta”) – odpowiednio: 99,9% i 98,9%.

Pod względem formy prawnej przedsiębiorstw największy odsetek (zarówno wśród przedsiębiorstw zlikwidowanych, jak i cenzurowanych) stanowiły osoby fizyczne prowadzące działalność gospodarczą, odpowiednio 97,6% i 89,1%.

Z punktu widzenia wielkości przedsiębiorstw dominowały mikroprzedsiębiorstwa (0–9 osób pracujących), w grupie jednostek zlikwidowanych i cenzurowanych stanowiły one odpowiednio 99,6% i 99,0%.

Spośród ogółu analizowanych przedsiębiorstw powstałych w województwie łódzkim, które uległy likwidacji w latach 2010–2015, największa część prowadziła działalność na terenie Łodzi (35,0%). Jednostki zlikwidowane w pozostałych powiatach województwa łódzkiego stanowiły od 0,75% (skierniewicki) do 6,76% (zgierski) ogółu zlikwidowanych przedsiębiorstw.

Analizując strukturę przedsiębiorstw zlikwidowanych w latach 2010-2015 według rodzaju prowadzonej działalności gospodarczej, należy zauważyć, że największy odsetek stanowiły jednostki z sekcji G – *Handel hurtowy i detaliczny...* (39,0%) oraz z sekcji F – *Budownictwo* (13,3%). Jednostki zlikwidowane w województwie łódzkim z pozostałych sekcji PKD stanowiły od 0,07% (B – *Górnictwo i wydobywanie*) do 8,4% (C – *Przetwórstwo przemysłowe*) ogółu zlikwidowanych przedsiębiorstw.

W przypadku rozważanych przedsiębiorstw mamy do czynienia z cenzurowaniem prawostronnym. Przedsiębiorstwa trafiały do analizy w różnym momencie w ciągu 6 lat (2190 dni) badań, ale zakończenie obserwacji przypadało na dzień 31 grudnia 2015 r. Badaną zmienną zależną był zatem czas trwania w dniach.

Do analizy czasu trwania przedsiębiorstw oraz identyfikacji zmiennych, które w sposób istotny związane są z ryzykiem zakończenia działalności, wykorzystano metodę Kapłana-Meiera, model proporcjonalnego hazardu Coxa oraz drzewa przeżycia.

² Rejestr REGON jest jedynym dostępnym źródłem danych o przedsiębiorstwach powstałych i zlikwidowanych – bez względu na formę prawną, gdyż z definicji zawiera dane o wszystkich podmiotach gospodarki narodowej. Mając na uwadze pewne niedoskonałości rejestru REGON, należy stwierdzić, że ustalenie dokładnej liczebności zdefiniowanej kohorty przedsiębiorstw nie było możliwe, stąd też uzyskane dane należy traktować jako dużą próbę badawczą [Markowicz 2012, s. 103].

Estymacja funkcji przeżycia metodą Kaplana-Meiera oraz modele regresji Coxa są znane, popularne i szczegółowo opisane w wielu pracach (por. np. [Stanisz (red.) 2005]).

Stosunkowo mało popularnym narzędziem w modelowaniu czasu trwania jest natomiast metoda rekurencyjnego podziału (por. np. [Cappelli, Zhang 2007; Bou-Hamad i in. 2011]). W metodzie rekurencyjnego podziału dokonywany jest podział zbioru obiektów na jednorodne podzbiory w oparciu o wartości zmiennych niezależnych, a następnie w każdym z węzłów końcowych szacowane są funkcje przeżycia, np. metodą Kaplana-Meiera lub budowane są modele regresji Coxa. Graficzną prezentacją metody są drzewa przeżycia.

Metoda rekurencyjnego podziału jest użytecznym narzędziem analizy czasu trwania. Wynika to z jej nieparametrycznego charakteru, braku wymagań co do rozkładu zmiennych i zależności między nimi. W prowadzonych analizach czasu trwania przedsiębiorstw wykorzystano dwa algorytmy budujące drzewa przeżycia: algorytm CTREE (*Conditional Inference Trees*), zaproponowany i szczegółowo opisany w pracy [Hothorn i in. 2006], oraz algorytm RPART, będący implementacją algorytmu CART [Breiman i in. 1984].

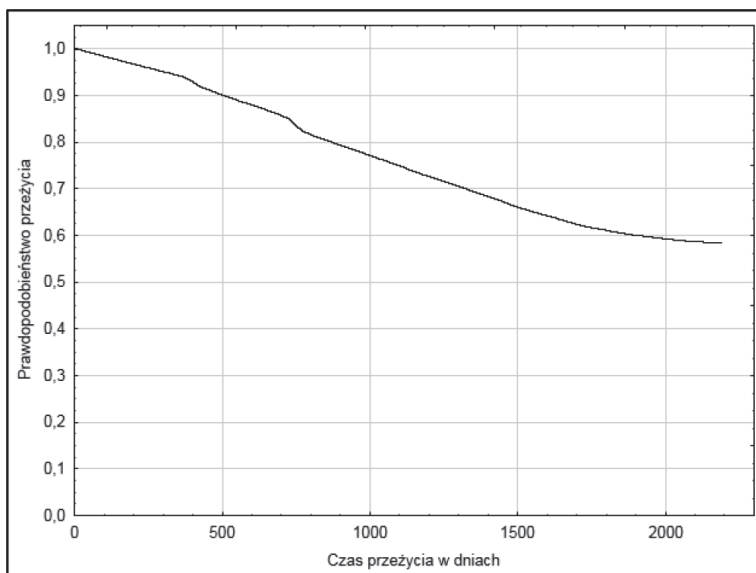
Wszystkie obliczenia wykonano z wykorzystaniem pakietu STATISTICA 12.5, SPSS 24 oraz w środowisku R v. 3.4.1 (pakiety: survival, rpart, party).

3. Wyniki badania czasu trwania przedsiębiorstw

Na rys. 1 przedstawiono krzywą przeżycia Kaplana-Meiera dla danych ogółem, bez wyodrębniania podgrup ze względu na charakterystyki przedsiębiorstw. Można zaobserwować ostrzejszy spadek krzywej przeżycia w ciągu pierwszych 2 lat (730 dni). Prawdopodobieństwo przeżycia 2190 dni (6 lat) jest równe 0,5845.

Aby oszacować parametry modelu Coxa, należało najpierw przekodować poszczególne zmienne jakościowe charakteryzujące przedsiębiorstwa, przyjmując jedną z kategorii za referencyjną. Uzyskany model był istotny statystycznie (LR $\chi^2 = 4815$; $df = 46$; $p < 0,001$). Nie było natomiast spełnione założenie o proporcjonalności hazardu ($p < 0,05$). Ze względu na ograniczoną objętość pracy zrezygnowano z prezentacji szczegółowych wyników, ograniczając się wyłącznie do opisu wyników istotnych statystycznie.

Ryzyko likwidacji przedsiębiorstwa w sposób istotny statystycznie zwiększa się w przypadku: osób fizycznych prowadzących działalność gospodarczą (współczynnik hazardu $HR = 2,185$; $p < 0,0001$), prowadzenia działalności w sekcji K wg PKD (*Działalność finansowa i ubezpieczeniowa*; $HR = 1,316$; $p = 0,0199$) oraz lokalizacji firmy w powiatach: kutnowskim ($HR = 1,147$; $p = 0,0021$), opoczyńskim ($HR = 1,123$; $p = 0,0208$), piotrkowskim ($HR = 1,112$; $p = 0,0277$), radomszczańskim ($HR = 1,166$; $p = 0,0002$), sieradzkim ($HR = 1,118$; $p = 0,0080$) i wieluńskim ($HR = 1,112$; $p = 0,0222$).

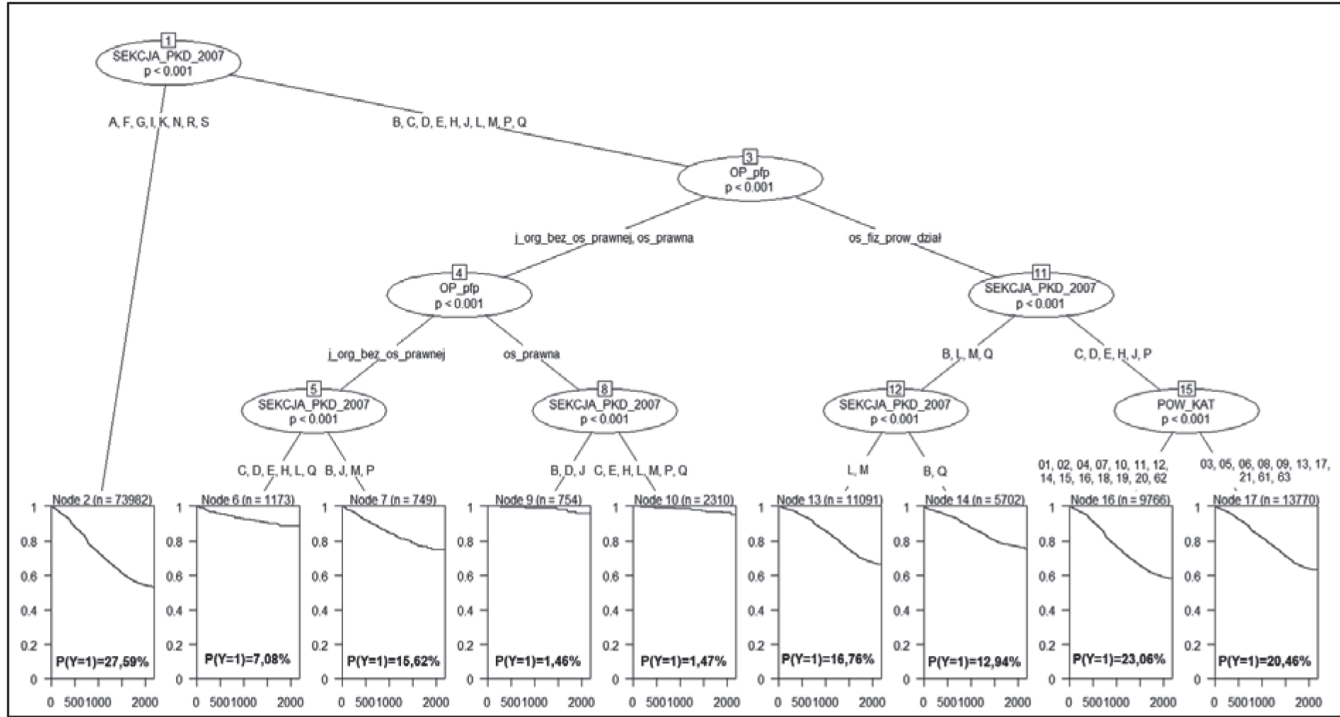


Rys. 1. Krzywa przeżycia Kaplana-Meiera dla ogółu przedsiębiorstw

Źródło: opracowanie własne w pakiecie STATISTICA.

Do czynników zmniejszających w sposób istotny statystycznie ryzyko likwidacji przedsiębiorstwa należą: prowadzenie działalności przez osoby prawne (HR = 0,0143; $p < 0,0001$), wielkość firmy mierzona liczbą pracowników 10–49 (HR = 0,608; $p < 0,0001$), prowadzenie działalności w następujących sekcjach wg PKD: B (*Górnictwo i wydobywanie* – HR = 0,590; $p = 0,0358$), C (*Przetwórstwo przemysłowe* – HR = 0,766; $p = 0,0223$), H (*Transport i gospodarka magazynowa* – HR = 0,768; $p = 0,0255$), J (*Informacja i komunikacja* – HR = 0,693; $p = 0,0024$), L (*Działalność związana z obsługą rynku nieruchomości* – HR = 0,615; $p = 0,0004$), M (*Działalność profesjonalna, naukowa i techniczna* – HR = 0,585; $p < 0,0001$), Q (*Opieka zdrowotna i pomoc społeczna* – HR = 0,454; $p < 0,0001$), forma własności: przedsiębiorstwa sektora prywatnego (własność mieszana; HR = 0,594; $p = 0,0062$) oraz lokalizacja siedziby przedsiębiorstwa w powiecie łódzkim wschodnim (HR = 0,864; $p = 0,0018$).

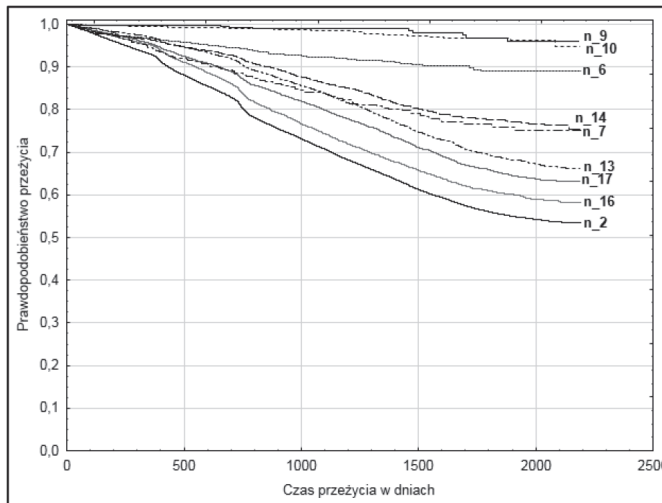
Drzewo przeżycia dla analizowanych przedsiębiorstw utworzone z wykorzystaniem algorytmu CTREE (pakiet party) przedstawiono na rys. 2. Drzewo ma 9 węzłów końcowych. Dla każdego z nich podano liczbę przedsiębiorstw, które znalazły się węzle w wyniku rekurencyjnego podziału, oraz odsetek przedsiębiorstw zlikwidowanych, a także narysowano krzywe przeżycia Kaplana-Meiera. W celu łatwiejszego porównania przeżycia firm z każdego węzła krzywe przeżycia Kaplana-Meiera zaprezentowano także na wspólnym rysunku (por. rys. 3).



Oznaczenia: SEKCJA_PKD_2007 – opis poszczególnych sekcji znajduje się w tekście; OP_pfp – forma prawna (jednostka organizacyjna bez osobowości prawnej, osoba prawna, osoba fizyczna prowadząca działalność gosp.); POW_KAT – oznaczenie powiatu (01: bełchatowski, 02: kutnowski, 03: łaski, 04: łączycki, 05: łowicki, 06: łódzki wschodni, 07: opoczyński, 08: pabianicki, 09: pajęczański, 10: piotrkowski, 11: poddębicki, 12: radomszczański, 13: rawski, 14: sieradzki, 15: skierniewicki, 16: tomaszowski, 17: wieluński, 18: wierzuszowski, 19: zduńskowolski, 20: zgierski, 21: brzeziński, 61: m. Łódź, 62: m. Piotrków Tryb., 63: m. Skierniewice).

Rys. 2. Drzewo przeżycia – algorytm CTREE

Źródło: opracowanie własne (pakiet party).



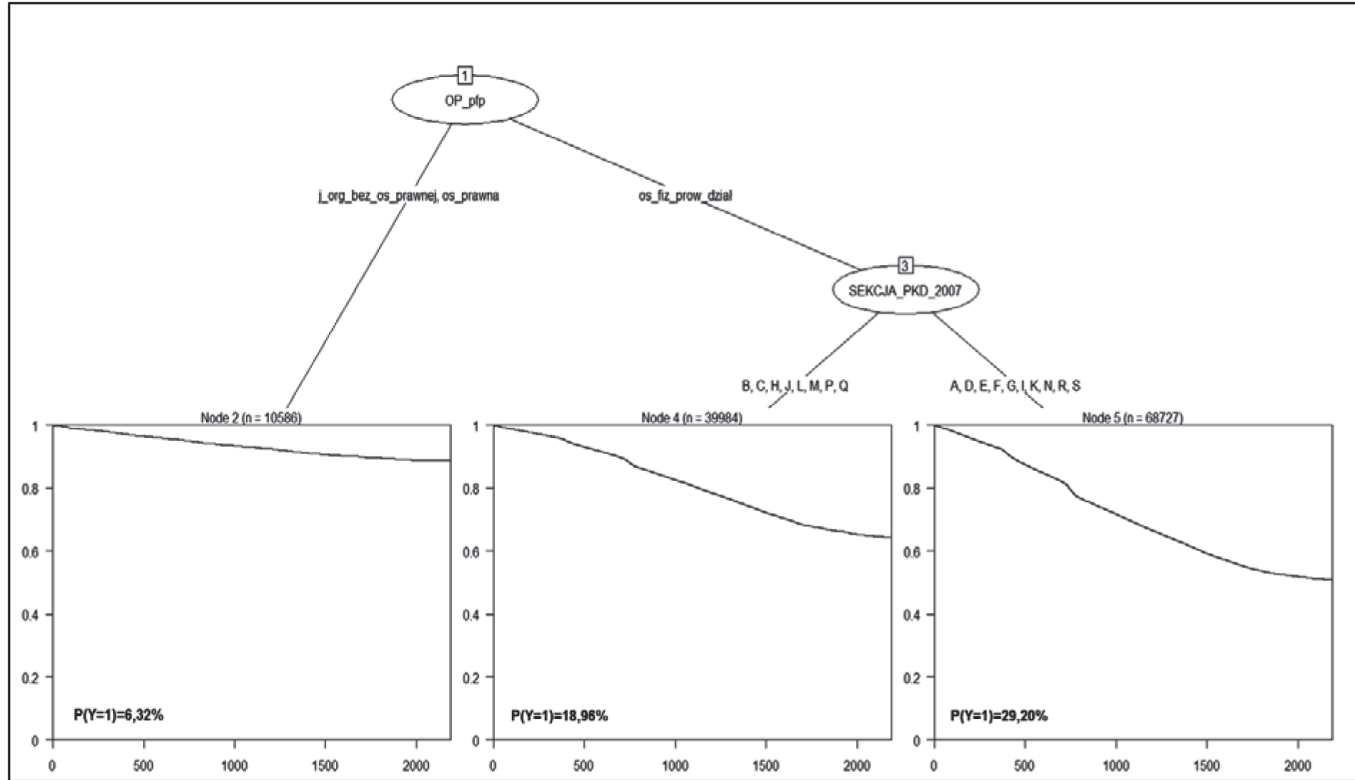
Rys. 3. Krzywe przeżycia Kaplana-Meiera w liściach drzewa CTREE

Źródło: opracowanie własne w pakiecie STATISTICA.

Jak można zauważyć, najwyżej położone są krzywe przeżycia dla węzłów nr 9 ($n = 754$) i nr 10 ($n = 2310$). W obu tych liściach znalazły się wyłącznie osoby prawne, zaklasyfikowane do następujących sekcji PKD: B (*Górnictwo i wydobywanie*), D (*Wytwarzanie i zaopatrywanie w energię elektryczną...*) i J (*Informacja i komunikacja*) – węzeł nr 9 lub C (*Przetwórstwo przemysłowe*), E (*Dostawa wody; gospodarowanie ściekami...*), H (*Transport i gospodarka magazynowa*), L (*Działalność związana z obsługą rynku nieruchomości*), M (*Działalność profesjonalna, naukowa i techniczna*), P (*Edukacja*), Q (*Opieka zdrowotna i pomoc społeczna*) – węzeł nr 10. Prawdopodobieństwo przeżycia 2190 dni jest równe odpowiednio: 0,9605 (węzeł nr 9) oraz 0,9487 (węzeł nr 10).

Najniżej położona krzywa Kaplana-Meiera znajduje się w węźle nr 2. W węźle tym znalazły się 73 982 przedsiębiorstwa. Są to jednostki zaklasyfikowane do sekcji PKD: A (*Rolnictwo, leśnictwo, łowiectwo i rybactwo*), F (*Budownictwo*), G (*Handel hurtowy i detaliczny...*), I (*Działalność związana z zakwaterowaniem...*), K (*Działalność finansowa i ubezpieczeniowa*), N (*Działalność w zakresie usług administrowania...*), R (*Działalność związana z kulturą, rozrywką i rekreacją*) i S (*Pozostała działalność usługowa*). Prawdopodobieństwo przeżycia 2190 dni wynosi 0,5343.

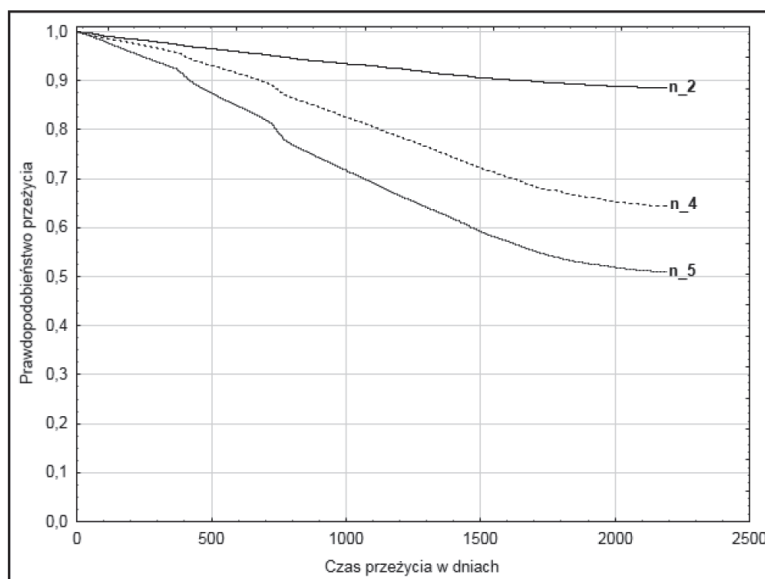
Testem log rank (Mantela-Coxa) odrzucono hipotezę zerową o równości funkcji przeżycia w wyróżnionych liściach ($p < 0,001$), zarówno dla analizy łącznej, jak i większości porównań parami. Brak istotnych statystycznie różnic między krzywymi przeżycia odnotowano dla trzech par porównań krzywych przeżycia: węzeł nr 9 vs węzeł nr 10 ($p = 0,442$), węzeł nr 7 vs węzeł nr 13 ($p = 0,233$) oraz węzeł nr 7 vs węzeł nr 14 ($p = 0,120$).



Oznaczenia: SEKCJA_PKD_2007 – opis poszczególnych sekcji znajduje się w tekście; OP_pfp – forma prawna (jednostka organizacyjna bez osobowości prawnej, osoba prawna, osoba fizyczna prowadząca działalność gosp.).

Rys. 4. Drzewo przeżycia – algorytm RPART

Źródło: opracowanie własne (pakiet rpart).



Rys. 5. Krzywe przeżycia Kaplana-Meiera w liściach drzewa RPART

Źródło: opracowanie własne w pakiecie STATISTICA.

Drzewo uzyskane z wykorzystaniem algorytmu RPART przedstawiono na rys. 4. Drzewo to ma 3 węzły końcowe. Dla każdego z nich podano liczbę przedsiębiorstw, które znalazły się węzle w wyniku rekurencyjnego podziału oraz odsetek przedsiębiorstw zlikwidowanych, a także narysowano krzywe przeżycia Kaplana-Meiera.

W celu łatwiejszego porównania przeżycia firm z każdego węzła krzywe przeżycia zaprezentowano na wspólnym rysunku (por. rys. 5).

Najwyżej położona jest krzywa przeżycia dla węzła nr 2 – w węzle tym znalazły się jednostki organizacyjne bez osobowości prawnej oraz osoby prawne ($n = 10586$). Prawdopodobieństwo przeżycia 2190 dni jest równe 0,8869.

Najniżej położoną krzywą Kaplana-Meiera uzyskano w węzle nr 5. W węzle tym znalazło się 68 727 przedsiębiorstw. Są to osoby fizyczne prowadzące działalność gospodarczą, sklasyfikowane w następujących sekcjach PKD: A (*Rolnictwo, leśnictwo...*), D (*Wytwarzanie i zaopatrywanie w energię...*), E (*Dostawa wody; gospodarowanie ściekami...*), F (*Budownictwo*), G (*Handel hurtowy i detaliczny...*), I (*Działalność związana z zakwaterowaniem...*), K (*Działalność finansowa i ubezpieczeniowa*), N (*Działalność w zakresie usług administrowania...*), R (*Działalność związana z kulturą, rozrywką i rekreacją*), S (*Pozostała działalność usługowa*). Prawdopodobieństwo przeżycia 2190 dni jest równe 0,5100.

Na podstawie wyników testu log rank (Mantela-Coxa) odrzucono hipotezę zerową o równości funkcji przeżycia w wyróżnionych liściach ($p < 0,001$), zarówno dla analizy łącznej, jak i poszczególnych porównań parami.

4. Zakończenie

Do podziału w węzłach drzew wykorzystane zostały 3 zmienne: forma prawna (OP_pfp), przynależność do sekcji PKD (SEKCJA_PKD_2007) oraz lokalizacja siedziby firmy (POWIAT). W żadnym z zastosowanych modeli drzew, w wyniku działania algorytmu, do podziału jednostek nie została wybrana wielkość przedsiębiorstwa mierzona liczbą osób pracujących i forma własności. Podobnie w modelu Coxa, istotne oszacowania parametrów dotyczyły przede wszystkim poszczególnych wariantów formy prawnej, wybranych sekcji PKD oraz niektórych powiatów.

Uzyskane wyniki (CTREE, RPART) są zgodne z wynikami innych badań – wskaźniki przetrwania przedsiębiorstw potwierdzają, że osoby fizyczne prowadzące działalność gospodarczą charakteryzują się niższymi krzywymi przeżycia [GUS 2012, 2013, 2014, 2015; Gurgul, Zajac 2016]. Jednocześnie jednostki zarejestrowane jako „osoby prawne” oraz prowadzące działalność gospodarczą w większości wskazanych sekcji PKD mają znacznie większe szanse na przetrwanie³. Analizując wyniki w węzłach 15 i 17 drzewa (algorytm CTREE), w których istotną rolę odgrywa lokalizacja przedsiębiorstwa, należy zauważyć, że powiaty z nieco wyżej położoną krzywą przeżycia (węzeł 17) to m.in. Łódź i okolice, powiat: brzeziński, łaski, łowicki, łódzki wschodni, pabianicki, pajęczański, rawski, wieluński – przyległe, „ułożone po przekątnej wewnątrz województwa łódzkiego”. Powiaty, w których ryzyko likwidacji firmy zwiększa się (wg modelu Coxa), to powiaty z reguły „przygraniczne” w województwie łódzkim, położone z dala od Łodzi. Jedynie lokalizacja siedziby przedsiębiorstwa w powiecie łódzkim wschodnim (centrum) stanowiła czynnik zmniejszający w sposób istotny statystycznie ryzyko jego likwidacji.

Metoda rekurencyjnego podziału może być stosowana jako uzupełnienie stosowanych zwykle metod – estymacji funkcji przeżycia metodą Kaplana-Meiera i modelu proporcjonalnego hazardu Coxa. Jej zaletą jest możliwość wyodrębnienia grup przedsiębiorstw podobnych pod względem czasu przeżycia i przejrzysta wizualizacja uzyskanych wyników w przypadku pojedynczych modeli. Poprawę stabilności utworzonych modeli drzew przeżycia można uzyskać wykorzystując modele zagregowane. Będzie to przedmiotem dalszych badań.

³ Działalność gospodarcza osób fizycznych, które nie zatrudniały pracowników na podstawie stosunku pracy (tzw. samozatrudnieni), często ulega likwidacji tuż po upływie 2 lat. Jest to okres, w którym przedsiębiorca, przy spełnieniu określonych ustawowo przesłanek, korzysta z preferencyjnych stawek składek na ubezpieczenie społeczne. Po tym okresie wysokość składek rośnie prawie 4,5-krotnie, może to być przyczyną utraty płynności finansowej przedsiębiorstwa i w konsekwencji likwidacji działalności. Inną przyczyną rezygnacji z prowadzenia działalności gospodarczej przez osoby fizyczne po 2 latach może być utrata dotacji przyznanej przez Urząd Pracy dla nowych przedsiębiorców. Warto także zauważyć, że przy likwidacji działalności gospodarczej prowadzonej indywidualnie (osoba fizyczna) nie jest wymagane przeprowadzenie postępowania likwidacyjnego przed wykreśleniem spółki z rejestru (osoba prawna).

Literatura

- Balicki A., 2006, *Analiza przeżycia i tablice wymieralności*, PWE, Warszawa.
- Bou-Hamad I., Larocque D., Ben-Ameur H., 2011, *A review of survival trees*, *Statistics Surveys*, vol. 5, s. 44–71.
- Breiman L., Friedman J.H., Olshen R.A., Stone C.J., 1984, *Classification and Regression Trees*, Chapman & Hall/CRC, Boca Raton, London–New York–Washington D.C.
- Cappelli C., Zhang H., 2007, *Survival trees*, [w:] *Statistical Methods for Biostatistics and Related Fields*, eds. W. Hardle, Y. Mori, P. Vieu, Springer, Berlin–Heidelberg, s. 167–179.
- Gurgul H., Zając P., 2016, *Modelowanie nieparametryczne czasu przeżycia przedsiębiorstw w Małopolsce*, *Wiadomości Statystyczne*, nr 12(667), s. 7–29.
- GUS, 2012 (2013, 2014, 2015), *Warunki powstania i działania oraz perspektywy rozwojowe polskich przedsiębiorstw powstałych w latach 2006-2010 (2007-2011, 2008-2012, 2009-2013)*, Informacje i opracowania statystyczne, Warszawa.
- Hothorn T., Hornik K., Zeileis A., 2006, *Unbiased recursive partitioning: A conditional inference framework*, *Journal of Computational and Graphical Statistics*, vol. 15(3), s. 651–674.
- Hozer J., Markowicz I., Stolorz B., 2008, *Zastosowanie metod analizy historii zdarzeń w badaniu czasu funkcjonowania firm*, Zapol, Szczecin.
- Jackowska B., 2015, *Analiza kohortowa czasu istnienia mikroprzedsiębiorstw w Gdańsku*, *Zarządzanie i Finanse. Journal of Management and Finance*, vol. 13, no. 4/2, s. 127–145.
- Markowicz I., 2012, *Statystyczna analiza żywotności firm*, *Rozprawy i Studia*, t. (CMIX) 835, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.
- Markowicz I., 2015, *Badanie kohort firm w analizie trwania*, *Ekonometria*, vol. 4(50), s. 105–125.
- Markowicz I., 2017a, *Analiza trwania firm w powiatach województwa zachodniopomorskiego*, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, nr 468, *Taksonomia 28, Klasyfikacja i analiza danych – teoria i zastosowania*, red. K. Jajuga, M. Walesiak, UE, Wrocław, s. 138–146.
- Markowicz I., 2017b, *Duration Model of Enterprises – Analysis of Territorial Groups*, *Institute of Economic Research Working Papers*, no. 71/2017, Toruń.
- Markowicz I., Stolorz B., 2006, *Analysis of the survival function of firms*, [w:] *Baltic Business Development: Regional Development, SME Management and Entrepreneurship*, red. J. Batóg, Szczecin.
- Markowicz I., Stolorz B., 2007, *Próba oceny ryzyka upadłości firmy*, *Zeszyty Naukowe Uniwersytetu Szczecińskiego. Finanse. Rynki finansowe. Ubezpieczenia*, nr 6, US, Szczecin, s. 199–206.
- Mikulec A., 2017a, *Analiza wskaźnikowa trwania przedsiębiorstw w województwie łódzkim*, *Wiadomości Statystyczne*, nr 11(678), s. 29–55.
- Mikulec A., 2017b, *Kohortowe tablice trwania przedsiębiorstw w województwie łódzkim – ujęcie kwartalne*, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, nr 468, *Taksonomia 28, Klasyfikacja i analiza danych – teoria i zastosowania*, red. K. Jajuga, M. Walesiak, UE, Wrocław, s. 147–160.
- Program badań statystycznych statystyki publicznej na rok 2015 – Rozporządzenie Rady Ministrów z dnia 27 sierpnia 2014 r.*, Dz.U. z 2014, poz. 1330 ze zm.
- Stanisz A. (red.), 2005, *Biostatystyka*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków.
- Szymański D., 2011, *Badanie żywotności nowo powstałych mikroprzedsiębiorstw w Polsce w latach 2002-2007*, rozprawa doktorska, <http://depotuw.ceon.pl/handle/item/177> (17.11.2017).
- Śmiech S., 2011, *Analiza przeżycia podmiotów gospodarczych w województwie małopolskim w latach 2002-2008*, *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie*, nr 876, UE, Kraków, s. 121–132.