

## Kilka uwag o modelowaniu umieralności osób starszych w Polsce

Słowa kluczowe: umieralność w starszym wieku, modelowanie umieralności

### 1. Wprowadzenie

Wraz ze starzeniem się populacji i wydłużaniem się życia ludzkiego coraz ważniejsze staje się dokładniejsze rozpoznanie prawidłowości rządzących umieralnością osób starszych. Wyróżnianie tzw. czwartej grupy wieku (oprócz dzieci, dorosłych, osób starszych) – osób sędziwych (tzw. starszych starszych – na przykład osób w wieku powyżej 80 lat) jest zatem w pełni uzasadnione. Cezura wieku tego rodzaju przyjmowana jest nie tylko w demografii, ale także w naukach społecznych, ekonomicznych, w biologii, gerontologii i in. (por. m.in. [Wachter, Finch, 1997; Grundy, 1997; Szukalski, 2003; Gazińska, 2006]). Zmiany w strukturze wieku i intensywne starzenie się populacji w różnych regionach świata niesie ze sobą poważne konsekwencje natury społecznej i ekonomicznej (por. m.in. [Clark, Spengler, 1980; Gazińska, Mojsiewicz, 2004; Grundy, 1997]). W związku z tym nie budzi zdziwienia fakt wzmożonego zainteresowania umieralnością osób w tej grupie wieku, a tym samym z możliwościami modelowania odpowiednich funkcji charakteryzujących wzorzec przeżycia dla tychże osób. Stosowanie modeli opisujących czas trwania życia ludzkiego dla osób sędziwych prowadzi najczęściej do przeszacowania umieralności, bowiem powyżej pewnego wieku zauważalne jest zwolnienie tempa wzrostu prawdopodobieństwa zgonów [Gazińska i in., 2006]. Przyjęcie hipotezy, o coraz wolniejszym tempie wzrostu prawdopodobieństwa zgonu osób sędziwych, uzasadnione jest za względu na to, że wraz z wiekiem nasila się selekcja, sprawiająca, że do wieku sędziwego dożywają osoby o najlepszym stanie zdrowia. Ulega także zmianie relacja pomiędzy czynnikami endo- i egzogenicznymi zgonów [Gazińska, Mojsiewicz, 2004; Horiuchi, Wilmoth, 1998; Szukalski, 2002].

Celem niniejszego artykułu jest zaprezentowanie wyników modelowania granicznego wieku, od którego rozpoczyna się zmiana tempa umieralności oraz przy jego pomocy weryfikacja hipotezy o coraz wolniejszym tempie wzrostu prawdopodobieństwa zgonu osób starszych w Polsce. Celem pomocniczym jest wskazanie przedziałów wieku o najwyższych przyrostach wartości intensywności umieralności.

Rozważania teoretyczne zweryfikowano z wykorzystaniem materiału statystycznego dotyczącego populacji Polski ogółem oraz w sześciu regionach z uwzględnieniem płci w 2004 roku.

### 2. Modele czasu trwania życia ludzkiego

Jeżeli przyjmiemy, że długość życia ludzkiego jest ciągłą zmienną losową, możemy zdefiniować funkcje użyteczne w analizie rozkładu czasu trwania życia, m.in. funkcje trwania życia i intensywności umieralności. Zmienną losową wyrażającą czas do wystąpienia zgonu osoby należącej do określonej populacji w momencie  $t=0$  oznaczmy przez  $T$  ( $T \geq 0$ ). Prawdopodobieństwo, że zgon wystąpi później niż w momencie  $t > 0$  lub, że osoba w czasie  $t > 0$  będzie należeć do zbiorowości (dożyje co najmniej do czasu  $t$ ), zapiszmy jako funkcję przeżycia (trwania życia)  $S(t) = P(T > t)$ ,  $S(0) = 1$ , która jest funkcją monotoniczną i nierosnącą oraz  $\lim_{t \rightarrow +\infty} S(t) = 0$ . Dystrybuanta zmiennej losowej  $T$ , będąca funkcją komplementarną do funkcji przeżycia  $F(t) = 1 - S(t)$ , określa model dożycia (dotrwania). Ważnym narzędziem w analizie przeżycia jest funkcja intensywności umieralności  $\mu(t)$ . Funkcja ta wyraża prawdopodobieństwo zgonu dla krótkich przyrostów czasu. Dla wyróżnionych funkcji (dystrybuantą, przeżycia i intensywności umieralności) oraz funkcji gęstości związanymi z czasem trwania

życia istnieją wzajemne zależności [Matłoka, 1997: 36].

Zmienna  $T$  jako zmienna losowa ma pewien rozkład. Jeżeli rozkład ten jest nieznan, model przeżycia jest nazywany nieparametryczny. W przeciwnym przypadku jest modelem parametrycznym. Inną klasę tworzą modele semiparametryczne, częściowo specyfikujące postać odpowiedniej funkcji. Jednymi z najstarszych nieparametrycznych modeli czasu trwania życia są tablice demograficzne, zwane tablicami trwania życia. Stanowią one grupę modeli nieparametrycznych, co oznacza, że nie znane są postaci analityczne funkcji charakteryzujących rozkład życia ludzkiego. W modelach tych szacujemy wartości funkcji dla dyskretnej zmiennej losowej  $T$ .

Okresowe tablice trwania życia przedstawiają hipotetyczny proces wymierania populacji (tzw. kohorty hipotetycznej) złożonej z różnych generacji na podstawie obserwacji ich umieralności w pewnym okresie (w naszym przypadku w 2004 roku).

Podstawowymi parametrami tablic trwania życia są prawdopodobieństwa przeżycia roku przez osobę w wieku  $x$  lat i prawdopodobieństwo zgonu w ciągu roku osoby w wieku  $x$  lat (a zatem odpowiednio funkcje  $S(t)$  i  $F(t)$ , które w tradycyjnym zapisie przyjmują postać  $p_x$  i  $q_x$ ). W badaniach wykorzystano dwa rodzaje okresowych tablic trwania życia, w których prawdopodobieństwa zgonu szacowano dla zbiorowości zmarłych II i III rodzaju (por. rys. 1).

W przypadku zbiorowości zmarłych II rodzaju prawdopodobieństwa zgonów szacowano jako:

$$q(t)^{\text{II}} = \frac{Z_{ml}(x, t-x-1) + Z_{st}(x+1, t-x-1)}{P(x, 1.01.t)}, \quad (1)$$

natomiast dla zbiorowości zmarłych III rodzaju:

$$q(x, t)^{\text{III}} = \frac{Z_{ml}(x, t-x-1) + Z_{st}(x, t-x)}{P(x, 1.01.t)}, \quad (2)$$

gdzie:  $x$  – wiek w liczbie ukończonych lat,  $t=2004$  – rok badania,  $Z_{st}$  – liczba zgonów starszych,  $Z_{ml}$  – liczba zgonów młodszych,  $P(x, 1.01.t)$  – liczba osób żyjących w wieku  $x$  lat w dniu 1 stycznia 2004 roku.

Zależność pomiędzy intensywnością umieralności i prawdopodobieństwem zgonów wyraża się wzorem:

$$q(x) = 1 - e^{-\int_x^{x+1} \mu(t) dt}. \quad (3)$$

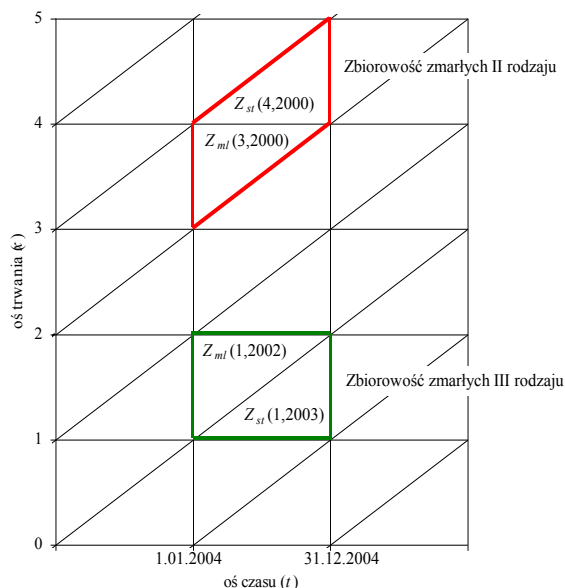
W praktyce szacując wartości empiryczne intensywności umieralności korzysta się z metody prostokątów całkowania numerycznego:

$$\int_x^{x+1} \mu(t) dt \cong \mu(x+1/2), \quad (4)$$

gdzie:  $\mu(x+1/2)$  – intensywność umieralności wyznaczona w połowie przedziału  $\langle x, x+1 \rangle$ .

Ze wzorów (3) i (4) wynika użyteczna zależność:

$$\mu(x+1/2) = -\ln(1-q(x)). \quad (5)$$



Rys. 1. Zbiorowości zmarłych II i III rodzaju w 2004 roku

Źródło: opracowanie własne

### 3. Modelowanie umieralności osób starszych

W wielu analizach statystycznych najwyższe (skrajne) obserwowane wartości mają istotne znaczenie. Kwantyfikacja ciężenia najwyższych kwantyli przyszłej długości trwania życia jest podstawą rozważań na temat granic funkcji gęstości tej zmiennej losowej [Kaufmann i in., 1997; Carey i in., 1992]. W modelowaniu intensywności umieralności osób starszych wykorzystuje się funkcje typu logistycznego lub inne funkcje o zbliżonym przebiegu [Gazińska i in., 2006]. Ważną własnością tych funkcji jest posiadanie punktu przegięcia, od którego wzrost tempa umieralności jest coraz wolniejszy.

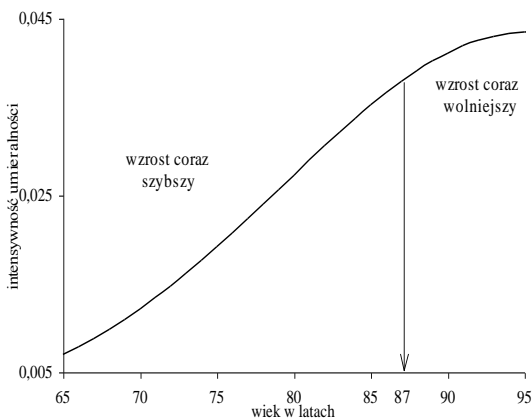
Przykładem takiej funkcji jest model autorstwa Coale'a i Kiskera, bazujący na założeniu, że stopa zmian współczynnika zgonów równa jest różnicy logarytmów współczynników zgonów osób starszych, co w konsekwencji prowadzi do funkcji kwadratowej postaci [Wilmoth, 1995]:

$$\ln[\mu(x)] = a + bx + cx^2, \text{ dla } c < 0. \quad (6)$$

Podstawową zaletą tego modelu jest prostota oraz fakt, że model ten można stosować w oparciu o dane okresowe, a nie wyłącznie dane kohortowe. Model ten został pozytywnie zweryfikowany na danych pochodzących z Japonii i Szwecji dla osób w wieku 85 lat i więcej [Gazińska i in., 2006; Wilmoth, 1995, cyt. za: Szukalski, 2002], a także dla danych z Polski i województwa zachodniopomorskiego [Gazińska, Mojsiewicz, 2004]. Przebieg przykładowej funkcji Coale'a-Kiskera zaprezentowano na rysunku 2. Na wykresie wyróżniono punkt przegięcia funkcji. Badania wskazują, że graniczny wiek, od którego następuje zmiana prawidłowości tempa umieralności oscyluje wokół 80-90 lat, w zależności od płci i regionu. Badania takie prowadzone były przy założeniu dopasowania intensywności umieralności do funkcji modelowej typu logistycznego lub wspomnianej funkcji Coale'a-Kiskera.

Rys. 2. Model Coale'a-Kiskera oszacowany dla intensywności umieralności kobiet powyżej 65 roku życia w województwie zachodniopomorskim w 2000 roku:

$$\mu(x) = \exp(-21,282 + 0,381 \cdot x - 0,002 \cdot x^2)$$



Źródło: opracowanie własne na podstawie [Gazińska, Mojsiewicz, 2004].

Należy jednak zwrócić uwagę na fakt, że wyniki modelowania dla danych z 2000 roku obarczone są błędem wynikającym z szacunków liczby ludności metodą bilansową na podstawie wyników Narodowego Spisu Powszechnego (NSP) z dnia 7.12.1988 roku oraz danych sprawozdawczości bieżącej, co więcej błędy te dotyczą zwłaszcza roczników mało licznych, a więc roczników starszych. W związku z tym należy liczyć

się z rozbieżnościami w modelowaniu dla danych z lat po 2002 roku, roku kolejnego NSP.

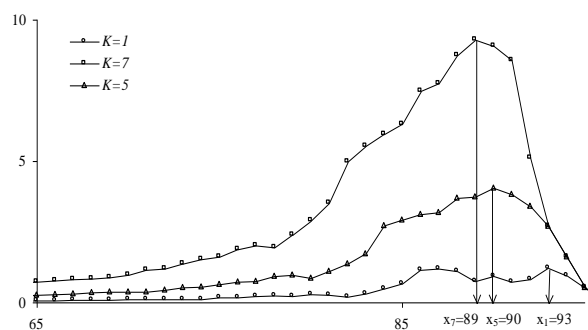
Szacowanie granicznego wieku, od którego rozpoczyna się zmiana tempa umieralności, proponuje się przeprowadzić niezależnie od dopasowania danych empirycznych do omawianych modeli. W tym celu konieczne jest ustalenie wieku, w którym intensywność umieralności posiada punkt przegięcia. W tym celu wykorzystano pierwszą pochodną intensywności umieralności obliczonej dla zbiorowości zmarłych II i III rodzaju na podstawie wzoru (5), która w tym punkcie osiąga maksimum. Pierwszą pochodną wyznaczono przy pomocy różniczkowania numerycznego.

Przy numerycznym różniczkowaniu funkcji zakłóconej losowo występuje zwielokrotnienie poziomu zakłócenia. W celu ograniczenia wpływu składnika losowego proponuje się następujący wzór, za [Mojsiewicz, Purczyński, 2005]:

$$\mu'(x) = \sum_{k=1}^K k \cdot (\mu(x+k) - \mu(x-k)), \text{ dla } K = 1, 2, \dots, 7. \quad (7)$$

We wzorze (7) zrezygnowano ze współczynnika skalującego, zależnego od  $K$ , ponieważ służy on tylko do wyznaczania wartości maksymalnej pochodnej. Przypadek  $K = 1$  odpowiada klasycznej regule różniczkowania numerycznego. Nie należy ograniczać się do przypadku  $K = 1$ , ponieważ obecność składnika losowego mogłaby spowodować wystąpienie maksimum. Zwiększając wartość  $K$  uzyskuje się wygładzenie krzywej odpowiadającej pochodnej  $\mu'(x)$ . Na rysunku 3 zaprezentowano wyniki wyznaczania wieku maksymalnej wartości różniczki dla zwiększanych wartości  $K$ .

Rys. 3. Różniczka numeryczna wyznaczona za (7) dla zbiorowości mężczyzn (zbiorowość zmarłych II rodzaju) w Polsce w 2004 roku

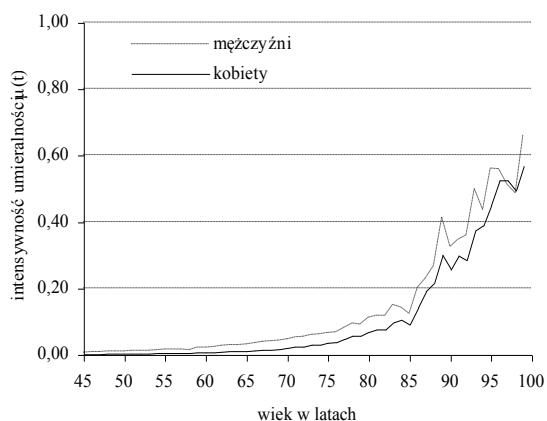


Źródło: obliczenia własne na podstawie BDR.

Na podstawie rysunku 3 można odczytać wartości  $x$  odpowiadające ekstremum: dla  $K = 1$ :  $x_1 = 93$ ; dla  $K = 7$ :  $x_7 = 89$ . Należy podkreślić, że przy analizie prawdopodobieństw zgonów bez stosowania wygładzenia szeregu pierwotnego, gdy przyjmujemy, że wygładzenie to dokonuje się przez różniczkę numeryczną przy wysokiej wartości parametru  $K$  zdarzają się oszacowania wykraczające poza empiryczny obszar zmienności wieku. Jako oszacowanie wieku zmiany tempa umieralności można przyjąć średnią arytmetyczną wartości  $x_1, \dots, x_7$  lub przyjąć wartość odpowiadającą maksymalnemu wygładzeniu szeregu przy nie przekraczaniu empirycznego obszaru zmienności.

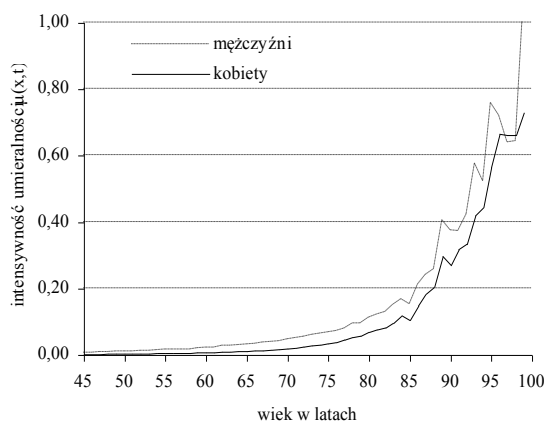
#### 4. Modelowanie umieralności osób starszych w Polsce

Na podstawie oszacowanych wartości prawdopodobieństwa zgonów dla ludności Polski oraz dla ludności w sześciu regionach dla 2004 roku obliczono wartości intensywności umieralności. Wartości intensywności umieralności dla zbiorowości zmarłych II i III rodzaju dla osób w wieku 45 lat i więcej według płci zamieszczono na rysunkach 4-17.



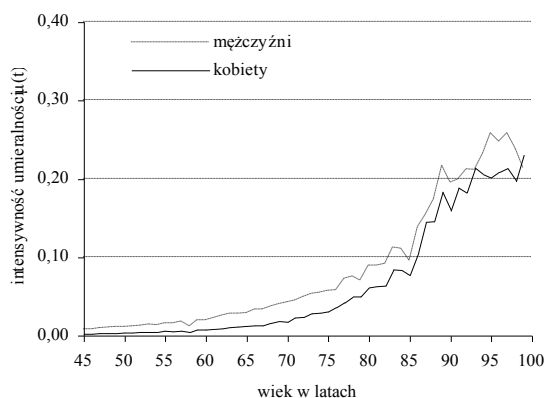
Rys. 4. Intensywności umieralności II rodzaju w Polsce w 2004 roku

Źródło: obliczenia własne na podstawie BDR.



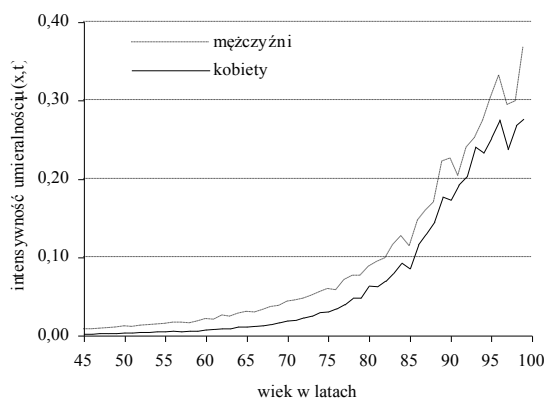
Rys. 5. Intensywności umieralności III rodzaju w Polsce w 2004 roku

Źródło: obliczenia własne na podstawie BDR.



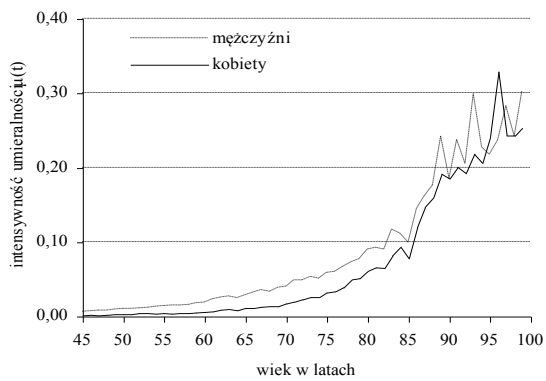
Rys. 6. Intensywności umieralności II rodzaju w regionie centralnym w 2004 roku

Źródło: obliczenia własne na podstawie BDR.

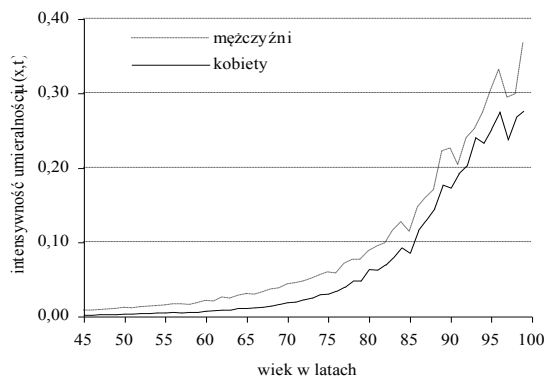


Rys. 7. Intensywności umieralności III rodzaju w regionie centralnym w 2004 roku

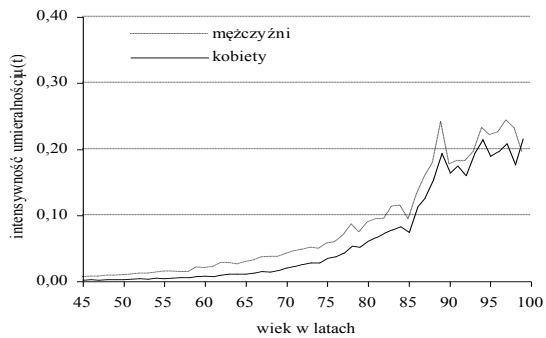
Źródło: obliczenia własne na podstawie BDR.



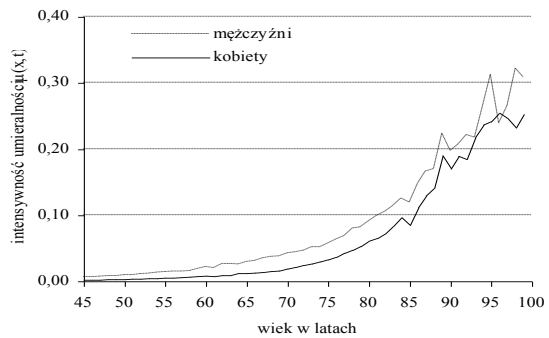
Rys. 8. Intensywności umieralności II rodzaju w regionie wschodnim w 2004 roku  
 Źródło: obliczenia własne na podstawie BDR.



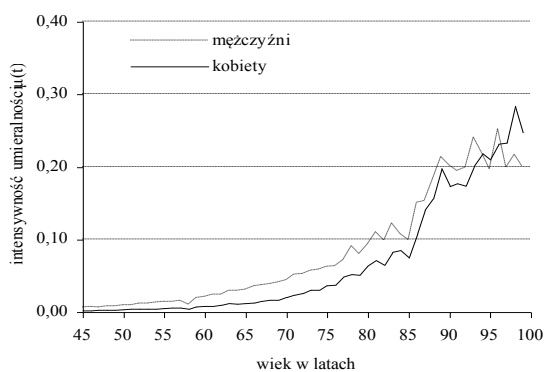
Rys. 9. Intensywności umieralności III rodzaju w regionie wschodnim w 2004 roku  
 Źródło: obliczenia własne na podstawie BDR.



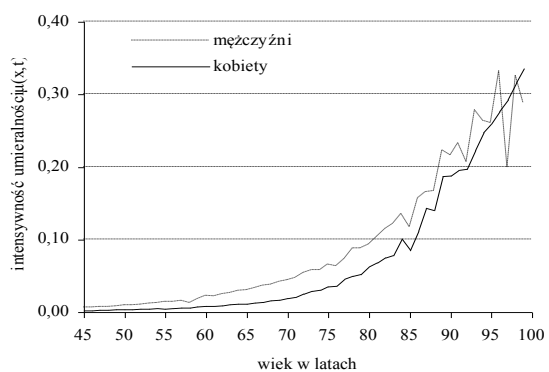
Rys. 10. Intensywności umieralności II rodzaju w regionie południowym w 2004 roku  
 Źródło: obliczenia własne na podstawie BDR.



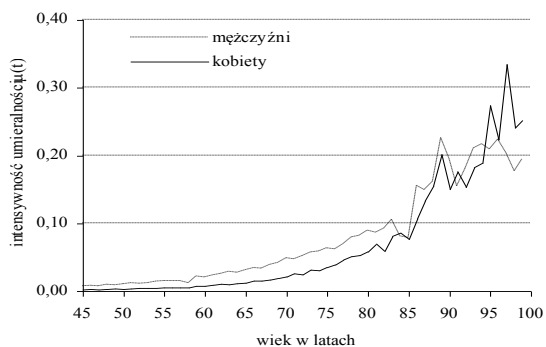
Rys. 11. Intensywności umieralności III rodzaju w regionie południowym w 2004 roku  
 Źródło: obliczenia własne na podstawie BDR.



Rys. 12. Intensywności umieralności II rodzaju w regionie północno-zachodnim w 2004 roku  
 Źródło: obliczenia własne na podstawie BDR.

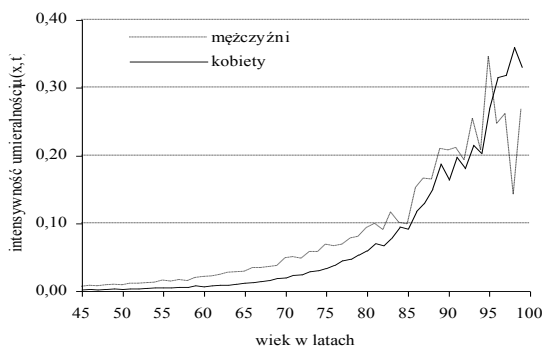


Rys. 13. Intensywności umieralności III rodzaju w regionie północno-zachodnim w 2004 roku  
 Źródło: obliczenia własne na podstawie BDR.



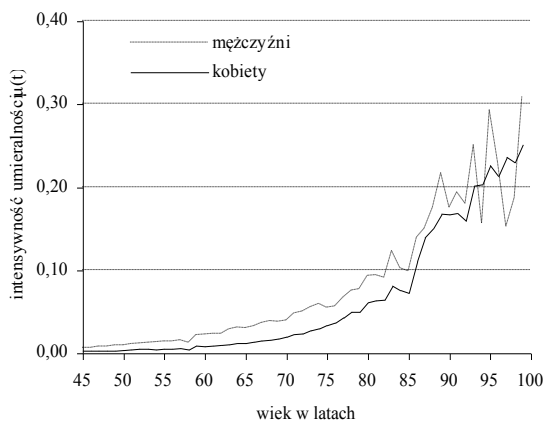
Rys. 14. Intensywności umieralności II rodzaju w regionie południowo-zachodnim w 2004 roku

Źródło: obliczenia własne na podstawie BDR.



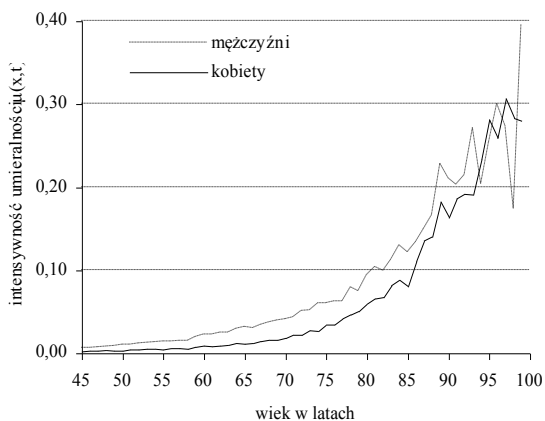
Rys. 15. Intensywności umieralności III rodzaju w regionie południowo-zachodnim w 2004 roku

Źródło: obliczenia własne na podstawie BDR.



Rys. 16. Intensywności umieralności II rodzaju w regionie północnym w 2004 roku

Źródło: obliczenia własne na podstawie BDR.



Rys. 17. Intensywności umieralności III rodzaju w regionie północnym w 2004 roku

Źródło: obliczenia własne na podstawie BDR.

Szacownie granicznego wieku osób starszych w Polsce oraz w sześciu regionach z uwzględnieniem płci przeprowadzono stosując numeryczne różniczkowanie

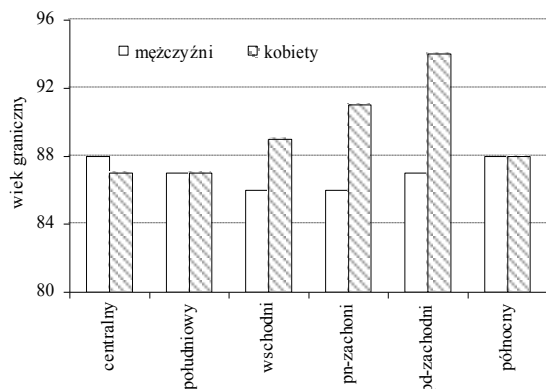
funkcji intensywności umieralności według wzoru (7). Wyniki szacowania zaprezentowano w tab. 1 oraz na rysunkach 18-19

Tab. 1 Wyniki szacowania punktu przęgięcia funkcji intensywności umieralności

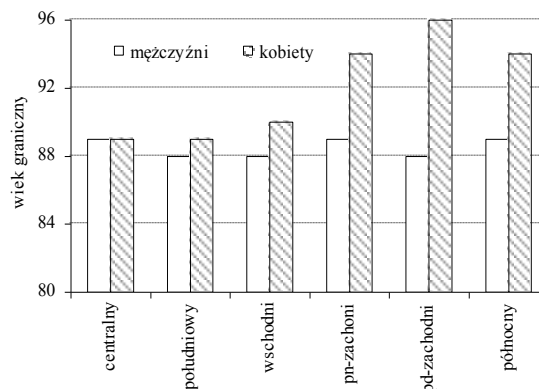
	Mężczyźni				Kobiety			
	zb. zmarłych II rodzaju		zb. zmarłych III rodzaju		zb. zmarłych II rodzaju		zb. zmarłych III rodzaju	
	$x_{sr}$	$x_{max}$	$x_{sr}$	$x_{max}$	$x_{sr}$	$x_{max}$	$x_{sr}$	$x_{max}$
Polska	89,4	89	91,6	90	94,0	94 <sup>*)</sup>	94,0	94 <sup>*)</sup>
Regiony:								
centralny	87,4	88	90,0	89	86,9	87	89,3	89
południowy	86,0	87	90,1	88	87,0	87	88,7	89
wschodni	87,6	86	89,0	88	89,9	89	92,9	90
pn-zachodni	87,1	86	90,9	89	87,3	91	88,8	94 <sup>*)</sup>
pd-zachodni	86,9	87	89,6	88	94,2	94 <sup>*)</sup>	95,5	96 <sup>**)</sup>
północny	88,1	88	88,9	89	87,0	88	94,2	94 <sup>*)</sup>

<sup>\*)</sup> dla  $K = 4$ , <sup>\*\*)</sup> dla  $K = 2$

Źródło: obliczenia własne.



Rys. 18. Wiek graniczny dla zbiorowości zmarłych II rodzaju w regionach w 2004 roku  
Źródło: tabela 1.



Rys. 19. Wiek graniczny dla zbiorowości zmarłych III rodzaju w regionach w 2004 roku  
Źródło: tabela 1.

#### 4. Podsumowanie

Problemy poruszane w niniejszym artykule nabierają znaczenia wraz ze zmianami umieralności oraz z procesami starzenia się populacji. Gwałtownie rośnie liczba osób w tzw. ekstremalnie starszych grupach wieku, czyli powyżej 90. roku życia. Zwiększa się w takiej sytuacji potrzeba dokładniejszego rozpoznania prawidłowości rządzących umieralności osób starszych.

Zaprezentowane modelowanie tzw. wieku granicznego, w którym zmienia się tempo intensywności umieralności w wykorzystaniem różniczki numerycznej może być alternatywnym podejściem modelowania umieralności osób starszych.

Wyniki modelowania wskazują wyraźnie na wyższe wartości wieku granicznego (dla obu płci we wszystkich regionach Polski) dla zbiorowości zmarłych III rodzaju. Dla obu zbiorowości zmarłych analizie poddano wiek graniczny odpowiadający maksymalnemu wygładzeniu szeregu przy nie przekroczeniu empirycznego obszaru zmienności otrzymując następujące wyniki:

- kobiety charakteryzują się wyższym wiekiem granicznym, ze wyjątkiem regionu centralnego,
- największe różnice między wiekiem granicznym kobiet i mężczyzn występują w regionach północno-zachodnim i południowo-zachodnim oraz północnym (dla zbiorowości zmarłych III rodzaju) – 5-8 lat,

- wiek graniczny mężczyzn nie wykazuje dużego zróżnicowania w regionach i można przyjąć, że oscyluje wokół 88 lat,
- wiek graniczny kobiet wykazuje zróżnicowanie w regionach – w regionach północnym, północno-zachodnim i południowo-zachodnim wiek graniczny kobiet osiąga wartości powyżej 90 lat, w pozostałych nie osiąga tego poziomu.

#### Literatura

- Carey J. R., Liedo P., Orozco D., Vaupel J. W., 1992, *Slowing of mortality rates at older ages in large medfly cohorts*, Science 258.
- Clark J.R., Spengler J.J., 1980, *The Economics of Individual and Population Ageing*, Cambridge University Press, Cambridge.
- Gazińska M., 2006, *Demograficzny próg starości*, [w:] Zeliaś A. *Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych*, Wyd. AE w Krakowie, Kraków.
- Gazińska M., Mojsiewicz M., 2004, *Modelowanie wartości górnych kwantyli rozkładu czasu trwania życia*, [w:] Tarczyński W. (red.) *Rynek kapitałowy, skuteczne inwestowanie*, Wyd. US, Szczecin.
- Gazińska M., Mojsiewicz M., Purczyński, 2006, *Modelowanie umieralności osób w wieku 80 lat i więcej*, „Studia Demograficzne” nr 1.
- Grundy E., 1997, *Demography and Gerontology: Mortality Trends Among the Oldest Old*, „Aging and Society”, 17.
- GUS (Główny Urząd Statystyczny), 2006, dane Banku Danych Regionalnych (BDR) dostępne na stronie internetowej [www.stat.gov.pl](http://www.stat.gov.pl) w dniu 15.01.1006.

- Horiuchi S., Wilmoth J. R., 1998, *Deceleration in the age pattern of mortality at older ages*, "Demography", vol. 35, nr 4.
- Kaninisto V., 1999, *The Advancing Frontier of Survival Life Table for Old Age*. Monographs on Population Aging, nr 3. Odense University Press.
- Kaufmann E., Reiss R. D., Thomas M., 2001, *About the Longevity of Humans*, [w:] *Statistical Analysis of Extreme Values*, Birkhaeuser Verlag, Basel-Boston-Berlin.
- Matłoka M., 1997, *Matematyka w ubezpieczeniach na życie*, Wyższa Szkoła Bankowa, Poznań.
- Mojsiewicz M., Purczyński J., 2005, *Modelowanie rozkładów szkód indywidualnych w warunkach moralnego hazardu*, [w:] *Inwestycje. Tendencje światowe a rynek polski*. Prace naukowe AE im. O. Langego we Wrocławiu, Wrocław.
- Szukalski P., 2002, *Parametryczne modele wymierania osób sędziwych*, „Wiadomości Ubezpieczeniowe”, nr 3/4.
- Szukalski P., 2003, *Tablice trwania życia dla osób długowiecznych*, „Wiadomości Ubezpieczeniowe”, nr 9/10.
- Wachter K.W., Finch C.E. (ed.), *Between Zeus and the Salmon*, 1997, National Academy Press, Washington.
- Wilmoth J. R., 1995, *Are mortality rates falling at extremely high ages? An investigation based on a model proposed by Coale and Kisker*, "Population Studies", vol. 45, nr 2.