

Agnieszka Rossa

Podstawy demometrii

Podstawy demometrii



WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

Agnieszka Rossa

Podstawy demometrii



**WYDAWNICTWO
UNIwersytetu
ŁÓDZKIEGO**

Łódź 2019

Agnieszka Rossa – Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny
Katedra Demografii, 90-214 Łódź, ul. Rewolucji 1905 roku nr 41/43

RECENZENT

Andrzej Ochocki

REDAKTOR INICJUJĄCY

Iwona Gos

KOREKTA JĘZYKOWA

Paweł M. Sobczak

SKŁAD I ŁAMANIE

Agnieszka Rossa

KOREKTA TECHNICZNA

Leonora Gralka

PROJEKT OKŁADKI

Katarzyna Turkowska

Zdjęcie wykorzystane na okładce: © Depositphotos.com/lamnee

Wydrukowano z gotowych materiałów dostarczonych do Wydawnictwa UŁ

© Copyright by Agnieszka Rossa, Łódź 2019

© Copyright for this edition by Uniwersytet Łódzki, Łódź 2019

Wydane przez Wydawnictwo Uniwersytetu Łódzkiego

Wydanie I. W.07706.16.0.S

Ark. druk. 9,375

ISBN 978-83-8142-374-8

e-ISBN 978-83-8142-375-5

<https://www.doi.org/10.18778/8142-374-8>

Wydawnictwo Uniwersytetu Łódzkiego

90-131 Łódź, ul. Lindleya 8

www.wydawnictwo.uni.lodz.pl

e-mail: ksiegarnia@uni.lodz.pl

tel. (42) 665 58 63

Spis treści

Przedmowa	7
Rozdział 1. ISTOTA I PRZEDMIOT DEMOGRAFII	9
1.1. Podstawowe definicje i pojęcia	9
1.2. Główne nurty demografii	10
1.3. Interdyscyplinarny charakter demografii	10
1.4. Źródła informacji w demografii	11
1.5. Pojęcie przestrzeni i czasu	12
1.6. Pojęcie czasu ekspozycji i osobołat życia	13
1.7. Siatki demograficzne	13
Ćwiczenie	20
Rozdział 2. WSPÓLCZYNNIKI DEMOGRAFICZNE	21
2.1. Wprowadzenie	21
2.2. Podział współczynników demograficznych	21
2.3. Współczynniki kohortowe	23
2.4. Współczynniki przekrojowo-kohortowe	26
2.5. Współczynniki przekrojowe	29
2.6. Wybrane współczynniki demograficzne	31
2.6.1. Ogólny współczynnik urodzeń (rodności)	31
2.6.2. Ogólny współczynnik płodności	32
2.6.3. Ogólny współczynnik zgonów	32
2.6.4. Współczynnik przyrostu naturalnego	33
2.6.5. Częstkowe współczynniki płodności	33
2.6.6. Częstkowe współczynniki zgonów	34
2.7. Przykłady	35
2.8. Prawdopodobieństwa zdarzeń demograficznych – ujęcie kohortowe i kohortowo-przekrojowe	49
2.9. Związek współczynników demograficznych i prawdopodobieństw zdarzeń	53
2.10. Standaryzacja i dekompozycja współczynników demograficz- nych	56
Ćwiczenia	65

Rozdział 3. TABLICE TRWANIA ŻYCIA	69
3.1. Wprowadzenie	69
3.2. Podstawowe charakterystyki TTŻ	69
3.3. Tablica kohortowa	72
3.4. Tablica przekrojowa	76
3.5. Strategie wyboru współczynników ${}_n a_x$	83
3.5.1. Model interpolacji liniowej	84
3.5.2. Model interpolacji wykładniczej	85
3.6. TTŻ w analizie czasu trwania zjawisk	90
3.7. Koncepcja populacji stacjonarnej	91
Ćwiczenia	95
Rozdział 4. ROZKŁADY CZASU TRWANIA ŻYCIA	101
4.1. Wprowadzenie	101
4.2. Niektóre charakterystyki funkcyjne i liczbowe	101
4.3. Intensywność zgonów a współczynnik zgonów	106
4.4. Relacja między ${}_y q_x$ a prawdopodobieństwami tablicowymi q_x	107
4.5. Prawa umieralności – ujęcie historyczne	110
Ćwiczenia	112
Rozdział 5. REPRODUKCJA LUDNOŚCI	113
5.1. Wprowadzenie	113
5.2. Wskaźnik dynamiki demograficznej	114
5.3. Współczynniki reprodukcji netto i brutto	114
5.4. Współczynnik dzietności	119
Ćwiczenie	121
Rozdział 6. WZROST LICZEBNY POPULACJI	123
6.1. Wprowadzenie	123
6.2. Równanie bilansowe	123
6.3. Intensywność wzrostu populacji	127
6.3.1. Roczna stopa wzrostu – przypadki szczególne	129
6.3.2. Liczba osobolat życia	133
6.4. Intensywność urodzeń, płodności, zgonów i migracji	134
6.5. Model Lotki ludności ustabilizowanej	136
6.5.1. Struktura ludności w modelu Lotki	139
6.5.2. Integralne równanie reprodukcji ludności	139
6.5.3. Oczekiwana długość generacji	141
6.5.4. Wyznaczanie współczynnika przyrostu naturalnego	142
6.5.5. Intensywność wzrostu a intensywność urodzeń, zgonów	144
6.5.6. Średnie trwanie życia a średni wiek w chwili zgonu	145
Ćwiczenia	147
Bibliografia	149

Przedmowa

Niniejsza książka zawiera omówienie podstawowych metod i modeli z zakresu demografii matematycznej, zwanej w skrócie demometrią.

Celem autorki było przybliżenie w możliwie przystępny i wizualnie czytelny sposób metod i narzędzi analizy demograficznej, poprzez zilustrowanie ich licznymi przykładami oraz zadaniami do samodzielnego rozwiązania. W opracowaniu pominięto elementy demografii opisowej, której głównym zadaniem jest opis liczbowy wielkości i struktur rzeczywistych zbiorowości ludzkich wraz zachodzącymi w nich procesami demograficznymi.

Materiał podzielony został na sześć rozdziałów. W rozdziale 1 przybliżone zostały istota i przedmiot demografii, jak również definicje podstawowych pojęć. Rozdział 2 zawiera omówienie ogólnych zasad konstrukcji współczynników demograficznych opartych na koncepcji siatki Lexisa. Rozdział 3 prezentuje podstawowe modele dyskretne stosowane w demografii, czyli tablice trwania życia, zwane także tablicami wymieralności. Alternatywą dla tego ujęcia są tzw. modele z czasem ciągłym. Niektóre z nich zostały przedstawione w rozdziale 4. Kolejny rozdział porusza podstawowe zagadnienia dotyczące reprodukcji ludności, natomiast rozdział ostatni kontynuuje tę tematykę w zakresie modeli wzrostu, w tym także modelu Lotki ludności ustabilizowanej.

Autorka pragnie podziękować Recenzentowi książki, Profesorowi Andrzejowi Ochockiemu, za cenne uwagi i sugestie, które pomogły poprawić, uzupełnić i uszczegółwić zagadnienia prezentowane w książce.

Rozdział 1

ISTOTA I PRZEDMIOT DEMOGRAFII

1.1. Podstawowe definicje i pojęcia

Demografia jest nauką o prawidłowościach rządzących rozwojem populacji, z uwzględnieniem konkretnych warunków społecznych i gospodarczych w danym okresie i na danym terytorium [Okólski 2004, Holzer 2006, Kędelski, Paradysz 2006].

Kluczowym w tej definicji jest pojęcie **populacji**. Określenie to powstało w XVI wieku, a jego rozumienie zmieniało się na przestrzeni lat i ma obecnie wiele znaczeń (np. populacja ludzi, firm, zwierząt). W demografii przez to pojęcie rozumieć będziemy populację ludzką, scharakteryzowaną pod względem:

- **podmiotowym** – zdefiniowanie jednostek wchodzących w skład badanej populacji, np. osoby lub grupy osób (gospodarstwa domowe, małżeństwa, rodziny, kohorty itp.) spełniające ustalone kryteria (społeczne, ekonomiczne, polityczne, etniczne),
- **przestrzennym** – ustalenie geograficznego umiejscowienia populacji,
- **czasowym** – określenie momentu lub okresu, w którym żyją jednostki populacji lub występują interesujące nas zdarzenia demograficzne.

Przedmiot badań demograficznych obejmuje stan i strukturę populacji wg cech społeczno-ekonomicznych, dynamikę zmian populacji, w tym zdarzenia i procesy demograficzne, ich zróżnicowanie, a także społeczno-ekonomiczne uwarunkowania i konsekwencje tych procesów.

Przez **procesy demograficzne** rozumieć będziemy zjawiska zachodzące w populacji, kształtowane przez natężenie i rozkład w czasie (kalendarz) **zdarzeń demograficznych**, do których należą przede wszystkim urodzenia, zgony, migracje oraz zmiany stanu cywilnego.

Obserwacji i analizie podlegają w szczególności:

- **zmienna zasobu**, tj. wielkość populacji w ustalonym momencie lub momentach czasu,
- **zmiennie strumienia**, tj. liczby zdarzeń demograficznych zrealizowanych w danym przedziale czasu, wpływających na zmiany w wielkości i strukturze populacji,
- **struktura populacji wg cech społeczno-ekonomicznych**, takich jak wiek, płeć, stan cywilny, wykształcenie, zawód,
- **struktura populacji wg cech geograficznych**, np. miejsca urodzenia, zamieszkania, pracy,
- **uwarunkowania i konsekwencje** społeczno-ekonomiczne procesów demograficznych.

1.2. Główne nurty demografii

Wyróżnić można następujące główne nurty demografii:

Demografia ogólna – zajmująca się teoretycznymi rozważaniami na temat prawidłowości rządzących rozwojem populacji, a także zagadnieniami związanymi z opisem i pomiarem zjawisk demograficznych.

Demografia opisowa – zajmująca się opisem liczbowym wielkości i struktur zbiorowości, a także zdarzeń i procesów demograficznych, z wykorzystaniem zarówno liczb bezwzględnych, jak i stosunkowych, np. wyrażonych za pomocą prawdopodobieństw określonych zdarzeń lub współczynników demograficznych.

Demografia matematyczna (demometria) – zajmująca się metodami pomiaru i opisu prawidłowości rządzących zbiorowościami i procesami demograficznymi oraz metodami ich modelowania i predykcji z wykorzystaniem aparatu matematyczno-statystycznego.

Poza tym podziałem istnieje też szereg specjalności szczegółowych, np. demografia społeczna, ekonomiczna, historyczna czy potencjalna.

1.3. Interdyscyplinarny charakter demografii

Obszary badawcze demografii zazębiają się z wielu innymi dyscyplinami, dla których sferą zainteresowań są populacje ludzkie.

Zagadnienia analizowane w ramach tych dyscyplin są często zbieżne z przedmiotem badań demograficznych. Do nich należą m.in.:

ekonomia – w zakresie badań dotyczących popytu, potencjału siły roboczej, uwarunkowań rozwoju systemów ekonomicznych,

socjologia – w obszarze badań nad rozwojem społeczeństw, zasadami formowania się rodzin, uwarunkowań rozwoju systemów społecznych,

geografia – w analizach nad migracjami i terytorialnym rozmieszczeniem ludności,

psychologia – w zakresie badań dotyczących postaw i zachowań ludzkich,

epidemiologia i nauki medyczne – w analizach śmiertelności, rozpowszechniania chorób, czynników wpływających na stan zdrowia i czas życia w zdrowiu.

1.4. Źródła informacji w demografii

Badania demograficzne opierają się na danych statystycznych dotyczących ludności, dlatego z demografią wiąże się tzw. **statystyka publiczna**.

W Polsce dane na temat stanu i struktury ludności w różnych przekrojach terytorialnych i czasowych są gromadzone i publikowane przez Główny Urząd Statystyczny (GUS) oraz wojewódzkie urzędy statystyczne. Bazy i rejestry tematyczne prowadzą także inne urzędy i organy administracji państwowej. Z punktu widzenia źródeł informacji, wyróżniamy (por. [Kurkiewicz (red.) 2010], s. 31–40):

- **powszechne spisy ludności** – badanie całkowite przeprowadzane mniej więcej co 10 lat (tj. w latach zakończonych na zero lub im bliskich),
- **rejestracja bieżąca** – rejestracja urodzeń, zgonów, zmian stanu cywilnego lub miejsca zamieszkania, prowadzona przez urzędy ewidencji ludności czy urzędy stanu cywilnego,
- **badania reprezentacyjne** – badania przeprowadzane w losowych próbach jednostek (osób, rodzin, gospodarstw domowych) i organizowane na użytek konkretnych celów. Przykładem jest Badanie Aktywności Ekonomicznej Ludności (BAEL) prowadzone przez GUS w losowej próbie gospodarstw domowych.

1.5. Pojęcie przestrzeni i czasu

Przestrzeń jest rozumiana w demografii nie tylko w kategoriach regionów geograficznych, ale również stanów i warstw społecznych. Przemieszczanie się osób pomiędzy regionami nazywane jest **mobilnością poziomą**, natomiast przemieszczanie się w obrębie warstw społecznych – **mobilnością pionową**.

Dalej stosować będziemy następujące pojęcia **czasu**:

1. **Czas własny** jednostek należących do kohort, generacji lub zbiorowości jednocześnie żyjących w danym czasie i na danym terytorium.

W przypadku osób indywidualnych czas własny zwykle utożsamia się z wiekiem, choć może być on rozumiany także np. jako czas pozostawania w wieku prokreacyjnym w przypadku kobiet, czas trwania związku w przypadku małżeństw, czas edukacji młodzieży szkolnej itp.

2. **Czas historyczny (kalendarzowy)**, w którym jest umiejscowiona dana kohorta, generacja lub zbiorowość ogółu żyjących na danym terytorium.

Bardziej szczegółowego wyjaśnienia wymagają użyte w tej klasyfikacji określenia: kohorta, generacja i zbiorowość ogółu żyjących.

Kohorta to zbiorowość jednostek wyodrębniona na podstawie zdarzeń demograficznych ustalonego typu, które wystąpiły w zadany okresie, np. dniu, miesiącu, kwartale, roku lub innym przedziale czasu i były doświadczeniem każdego z członków tej zbiorowości. Mówimy wtedy o kohortach dziennych, miesięcznych, kwartalnych, rocznych itp.

Kohortę mogą tworzyć np. małżeństwa zawarte w danym roku kalendarzowym, kobiety rodzące w tym samym dniu, dzieci rozpoczynające naukę w ustalonym roku szkolnym, absolwenci podejmujący pierwszą pracę w określonym miesiącu itp.

Szczególny przypadek stanowi **kohorta urodzeniowa**, którą tworzą osoby urodzone w tym samym okresie kalendarzowym. Gdy rozważanym okresem jest rok, wówczas taką kohortę nazywać będziemy **generacją**.

Zbiorowość ogółu żyjących tworzą jednostki z różnych generacji żyjących jednocześnie w danym czasie. Prawidłowości zaobserwowane w takiej zbiorowości odnosi się niekiedy do tzw. **kohorty hipotetycz-**

nej, o której zakłada się, że intensywność (natężenie) określonych zdarzeń demograficznych jest zgodna z natężeniem zdarzeń odnotowanym w badanym okresie we wszystkich generacjach wchodzących w skład tej zbiorowości.

1.6. Pojęcie czasu ekspozycji i osobolat życia

W analizie demograficznej ważną rolę pełni pojęcie **czasu ekspozycji** (*person-years of exposure*). Rozumie się go jako czas, przez który dana jednostka jest narażona na ryzyko wystąpienia określonego zdarzenia demograficznego. Suma czasów ekspozycji dla wszystkich jednostek zbiorowości będzie nazywana dalej łącznym czasem ekspozycji.

Przykładowo, kobieta wchodząca w wiek rozrodczy ma szansę urodzenia dziecka. Mówimy wówczas, że podlega „ryzyku” doświadczenia tego zdarzenia demograficznego, a ekspozycja na ryzyko rozpoczyna się w momencie osiągnięcia wieku prokreacyjnego. W przypadku innego zdarzenia, jakim jest np. zgon, na ryzyko narażone są wszystkie jednostki zbiorowości, a czas ekspozycji mierzony jest od momentu urodzenia się poszczególnych osób.

Czas ekspozycji zastępowany jest w niektórych przypadkach pojęciem **osobolat życia** (*person-years lived*). Pojęcia tego używać będziemy w sytuacji, gdy nie wszystkie jednostki zbiorowości narażone są na ryzyko wystąpienia danego zdarzenia [Preston et al. 2001].

1.7. Siatki demograficzne

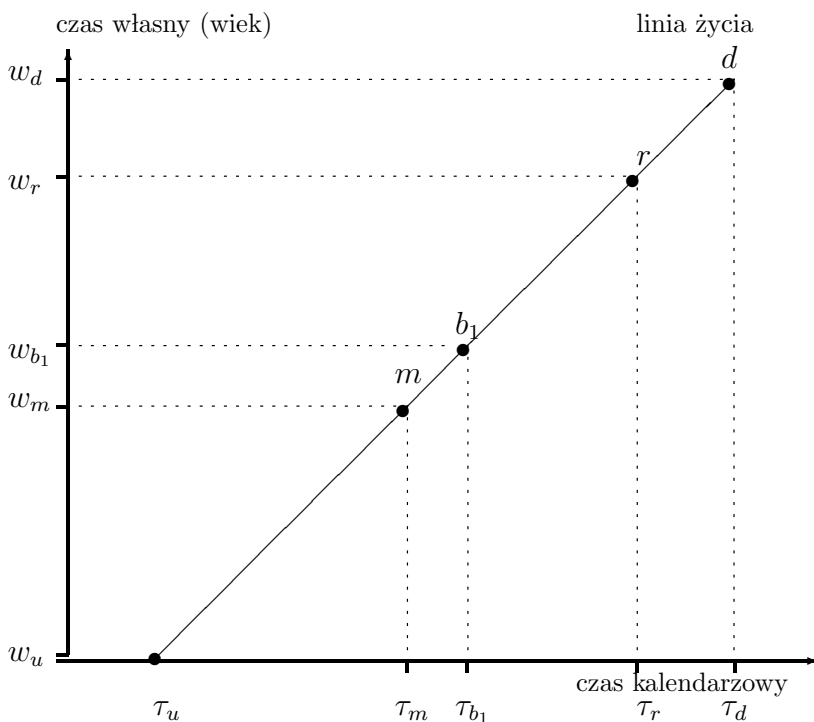
Częstą praktyką prezentacji kohort czy generacji jest umiejscowienie ich w dwuwymiarowym układzie współrzędnych, w którym oś pionowa reprezentuje czas własny kohorty (generacji), a oś pozioma – czas kalendarzowy. Taki rodzaj prezentacji nazywamy **siatką demograficzną**. Począwszy od lat 60. XX wieku, bardzo popularny stał się jej francuski wariant, będący modyfikacją tradycyjnej siatki zaproponowanej przez W. Lexisa w 1875 roku.

W rozważaniach i analizach opartych na siatkach demograficznych będziemy zakładać dalej, że mamy do czynienia z kohortami (generacjami), w których przyrost lub ubytek liczebny spowodowany jest ruchem naturalnym (urodzenia, zgony), natomiast nie występuje ruch

wędrowski (odpływ lub napływ migracyjny). Tego rodzaju zbiorowości nazywamy **zamkniętymi**.

Przykład 1. Przedstawimy graficzną prezentację losów pewnej kobiety w postaci tzw. **linii życia**, wykreślonej na rysunku 1 za pomocą linii prostej, nachylonej pod kątem 45° do osi poziomej.

Założmy, że dysponujemy danymi na temat momentów czasowych wystąpienia następujących zdarzeń, będących udziałem tej kobiety: u – narodziny, m – zawarcie małżeństwa, b_i – urodzenie i -tego dziecka, r – rozpad małżeństwa (tj. rozwód lub śmierć małżonka), d – śmierć.



Rysunek 1. Przykładowa linia życia w układzie współrzędnych

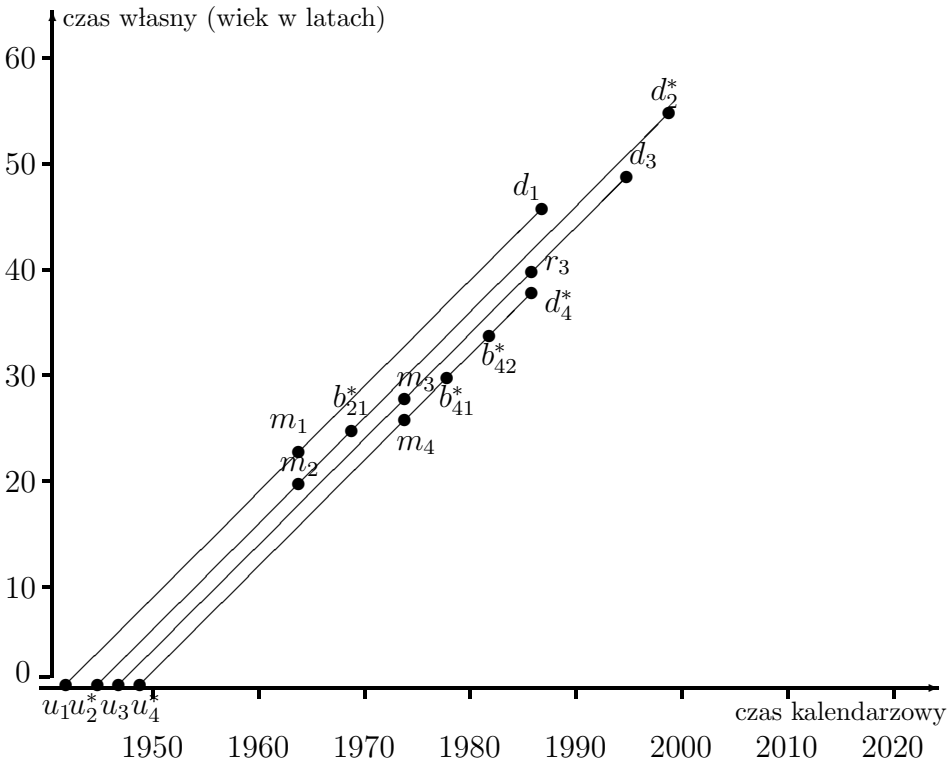
Obok linii życia, na rysunku 1 zaznaczono umiejscowienie w czasie własnym i kalendarzowym poszczególnych zdarzeń w dwuwymiarowym układzie współrzędnych.

Współrzędne na osi pionowej reprezentują wiek kobiety w chwili wystąpienia poszczególnych zdarzeń, natomiast współrzędne na osi poziomej – momenty kalendarzowe tych zdarzeń. Współrzędne te oznaczone zostały symbolami odpowiednio:

$w_u, w_m, w_{b_i}, w_r, w_d$ – momenty czasu własnego zaznaczone na osi pionowej, reprezentujące wiek w chwili wystąpienia zdarzeń u, m, b_i, r, d ,
 $\tau_u, \tau_m, \tau_{b_i}, \tau_r, \tau_d$ – momenty czasu kalendarzowego zaznaczone na osi poziomej, reprezentujące daty kalendarzowe zdarzeń u, m, b_i, r, d .

Przedstawiony schemat ilustruje zdarzenia demograficzne w odniesieniu do pojedynczej osoby. W demografii przedmiotem zainteresowania zwykle nie są pojedyncze jednostki, ale duże zbiorowości jednostek. Jednakże w większych zbiorowościach nie ma na ogół możliwości odnotowania dokładnych momentów pojedynczych zdarzeń. Zamiast tego rozpatruje się liczby zagregowanych zdarzeń, które wystąpiły w zadanych przedziałach czasu kalendarzowego lub własnego. Liczby te umieszcza się wewnątrz diagramu podzielonego na jednostkowe kwadraty, tj. pola o wymiarach 1×1 (jednostka czas własnego \times jednostka czasu kalendarzowego). Zagadnienie to wyjaśnimy w przykładzie 2.

Przykład 2. Rozważmy linie życia czterech osób urodzonych w latach 1940 – 1950. W tym przykładzie będą to dwa małżeństwa (rysunek 2).



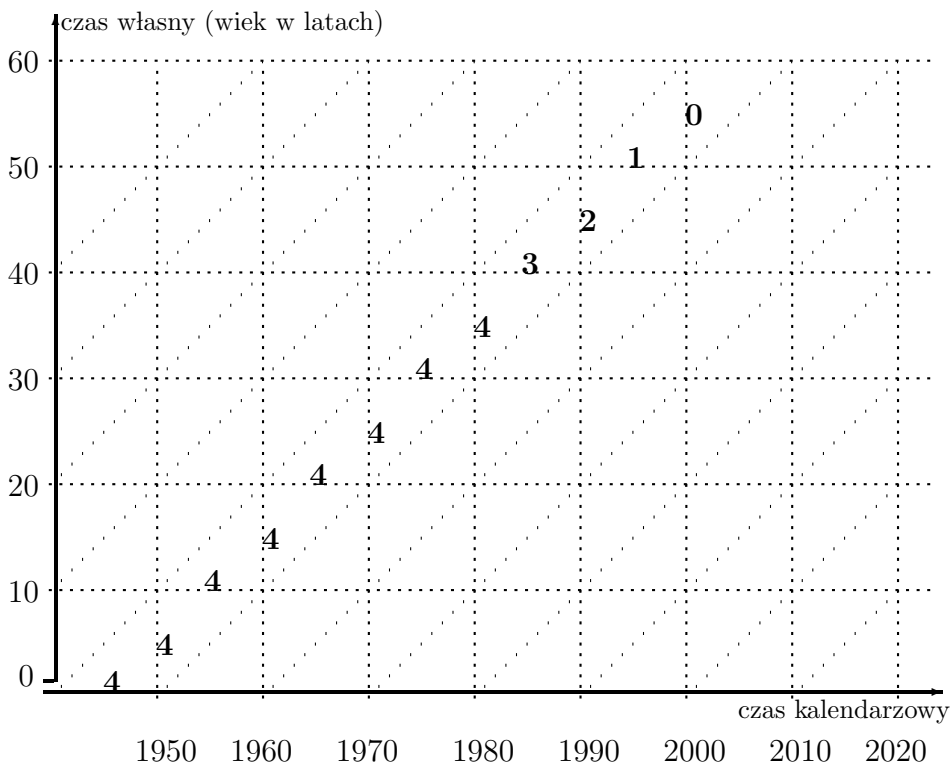
Rysunek 2. Przykładowe cztery linie życia

Podobnie jak w przykładzie 1, zdarzenia demograficzne oznaczone zostały etykietami $u_k, m_k, b_{ki}, r_k, d_k$, przy czym dodatkowy indeks k oznacza numer jednostki. Zdarzenia dotyczące kobiet wyróżnione zostały dodatkowo gwiazdką*.

Rysunek 3 przedstawia fragment siatki demograficznej z poziomymi oraz pionowymi liniami układu współrzędnych, wykreślonymi w jednostkowych odstępach czasu (tu co 10 lat) oraz z liniami ukośnymi, dzielącymi każdy kwadrat jednostkowy na dwa trójkąty – górny i dolny, zwane **trójkątami Lexisa**.

Przy wybranych odcinkach poziomych i pionowych siatki umieszczone zostały liczby osób z analizowanej czterosobowej kohorty, które dożyły kolejnych momentów czasu własnego i kalendarzowego. Są to tzw. liczby dożywających I i II rodzaju ilustrujące proces wymierania kohorty.

Wewnątrz trójkątów Lexisa można umieścić dodatkowo liczby zdarzeń określonego rodzaju, np. liczby zgonów. W tym przykładzie zostały one pominięte dla zachowania lepszej przejrzystości wykresu.

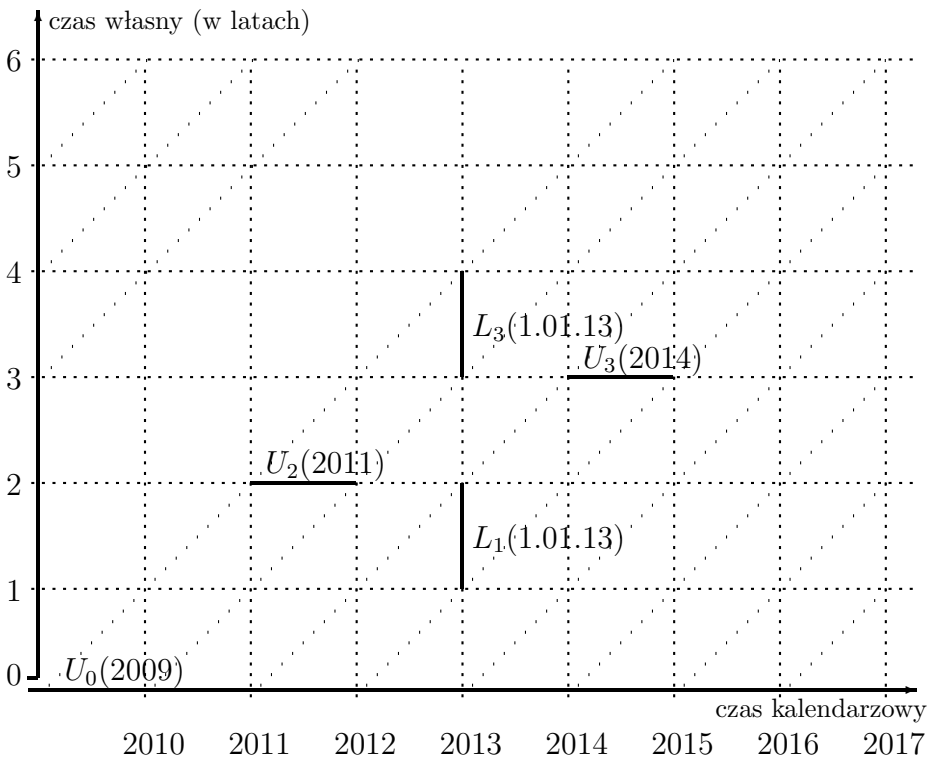


Rysunek 3. Fragment siatki demograficznej i liczby jednostek dożywających (dane na podstawie rysunku 2)

Wyjaśnimy bardziej szczegółowo pojęcie **dożywających I i II rodzaju**, użyte w przykładzie 2. W tym celu rozważmy podział czasu własnego na roczne przedziały $[x, x + 1)$. Niech t oznacza wybrany rok kalendarzowy, czyli przedział czasu od 1 stycznia do 31 grudnia roku t , natomiast τ niech będzie ustalonym momentem w ciągu danego roku, określonym z dokładnością do jednego dnia. Na przykład, dzień 1 stycznia roku t oznaczymy symbolem $\tau = 1.01.t$.

Zbiorowość dożywających I rodzaju – zbiór jednostek, które osiągnęły czas własny x (pełnych lat) w ciągu roku t . Liczbę takich jednostek oznaczać będziemy przez $U_x(t)$. W niektórych sytuacjach zapis może być skrócony do U_x .

Zbiorowość dożywających II rodzaju – zbiór jednostek, które w ustalonym momencie τ czasu kalendarzowego były w przedziale czasu własnego $[x, x + 1)$ lat, tj. czas własny (np. wiek) wynosił x pełnych lat. Liczbę takich jednostek oznaczać będziemy symbolem $L_x(\tau)$. W szczególnych przypadkach zapis można skrócić do $L(\tau)$ lub L_x .



Rysunek 4. Fragment siatki demograficznej i liczby dożywających I i II rodzaju

Przykładowe oznaczenia liczb dożywających I i II rodzaju umieszczono na rysunku 4 przy wybranych poziomych i pionowych odcinkach siatki.

Poza liczbami dożywających I i II rodzaju, w siatce Lexisa umieszcza się także liczby zdarzeń demograficznych. Zdarzenia przypisane do górnych i dolnych trójkątów Lexisa określane są terminami odpowiednio **zdarzeń młodszych** i **starszych**.

Liczby zdarzeń młodszych i starszych przyporządkowane do czasu własnego x (pełnych lat) w roku kalendarzowym t oznaczać będziemy symbolami odpowiednio $Z_x^{ml}(t)$ oraz $Z_x^{st}(t)$.

Ze względu na sposoby łączenia dwóch sąsiadujących ze sobą trójkątów Lexisa, wyróżniamy trzy podstawowe ich konfiguracje, przyjmujące postać równoległoboków poziomych, pionowych lub kwadratów [por. rysunek 5]. Korespondują one z trzema rodzajami zbiorów zdarzeń demograficznych.

Zbiór zdarzeń I rodzaju – reprezentowany przez równoległobok poziomy siatki Lexisa. Zawiera zdarzenia młodsze i starsze odnotowane w latach odpowiednio $t + 1$ oraz t we wspólnym przedziale czasu własnego $[x, x + 1)$. Zauważymy, że zdarzenia te dotyczą jednej kohorty powstałej (lub urodzonej) w roku $s = t - x$, którą nazwiemy w skrócie kohortą s . Sumę tych zdarzeń identyfikujemy jednoznacznie na podstawie czasu własnego i numeru kohorty. Oznaczając ją jako $Z_x^{koh.s}$, mamy

$$Z_x^{koh.s} = Z_x^{ml}(t + 1) + Z_x^{st}(t). \quad (1.1)$$

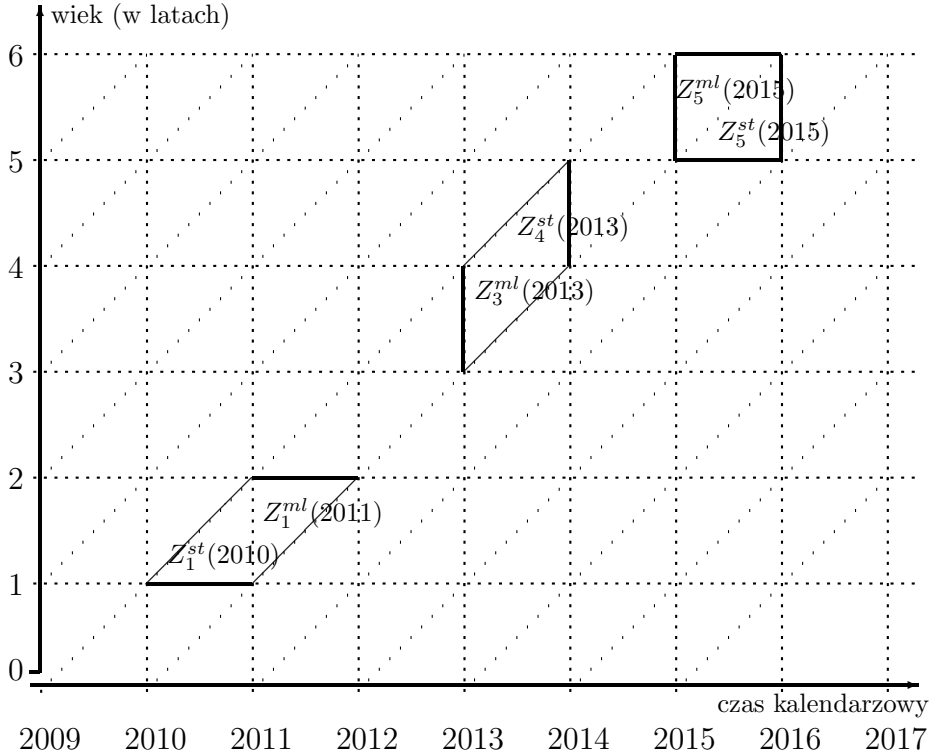
Zbiór zdarzeń II rodzaju – koresponduje z równoległobokiem pionowym siatki Lexisa. Zawiera zdarzenia młodsze i starsze w przedziałach czasu własnego odpowiednio $[x - 1, x)$ i $[x, x + 1)$ odnotowane we wspólnym roku kalendarzowym t . Zdarzenia te identyfikujemy za pomocą oznaczenia kohorty $s = t - x$ i okresu t . Suma $Z^{koh.s}(t)$ wynosi

$$Z^{koh.s}(t) = Z_{x-1}^{ml}(t) + Z_x^{st}(t). \quad (1.2)$$

Zbiór zdarzeń III rodzaju – wyrażony za pomocą kwadratu siatki Lexisa. Zawiera zdarzenia młodsze i starsze odnotowane we wspólnym przedziale czasu własnego $[x, x + 1)$ i roku t . Zdarzenia te dotyczą jednostek pochodzących z dwóch sąsiednich kohort $s = t - x - 1$ oraz $s + 1 = t - x$. Do oznaczenia ich sumy użyjemy symbolu $Z_x(t)$, ponieważ identyfikujemy ją na podstawie indeksów x, t . Mamy

$$Z_x(t) = Z_x^{ml}(t) + Z_x^{st}(t). \quad (1.3)$$

Przykładową ilustrację graficzną liczb zdarzeń I, II i III rodzaju w wybranych przedziałach czasu własnego i kalendarzowego przedstawia rysunek 5.



Legenda: zdarzenia I rodzaju – reprezentowane przez równoległobok poziomy, zdarzenia II rodzaju – reprezentowane przez równoległobok pionowy, zdarzenia III rodzaju – reprezentowane przez kwadrat Lexisa (zdarzenia młodsze, starsze – odpowiednio górny i dolny trójkąt Lexisa).

Rysunek 5. Fragment siatki demograficznej, liczby zdarzeń I, II, III rodzaju oraz zdarzeń młodszych i starszych

Na rysunku zaznaczono równoległobok poziomy, pionowy oraz kwadrat złożone z dwóch przylegających trójkątów Lexisa. Każdy z trójkątów przedstawia w sposób graficzny zbiory zdarzeń młodszych lub starszych. Ich liczby oznaczone są umownie symbolami $Z_x^{ml}(t)$, $Z_x^{st}(t)$, gdzie x jest wiekiem w ukończonych latach, a t – rokiem kalendarzowym. Trzy różne złożenia trójkątów Lexisa stanowią reprezentację zbiorów zdarzeń I, II, III rodzaju odnotowanych w odpowiednich przedziałach czasu własnego i kalendarzowego. Dwa pierwsze zbiory dotyczą osób urodzonych w roku 2009, a ostatni – generacji z lat 2009 i 2010.

Ćwiczenie

Dany jest fragment siatki demograficznej (rysunek 6), przedstawiający umowne liczby kobiet (w tys.) z dwóch generacji, dożywających wieku 20, 21, ..., 25 lat (liczby pogrubione na odcinkach poziomych siatki), a także liczby urodzonych dzieci (w tys.) w podziale wg wieku matek (liczby pochylone, wewnątrz trójkątów Lexisa).

- Ile dzieci urodziły kobiety należące do generacji z roku 1990, będąc w grupie wieku:
 - 20 ukończonych lat?
 - 21 ukończonych lat?
- Podać liczbę dzieci urodzonych w obu generacjach kobiet w roku:
 - 2010,
 - 2011,
 - 2012.
- Obliczyć liczbę zgonów wśród kobiet w wieku 23 ukończonych lat:
 - w obu generacjach,
 - w generacji urodzonych w 1990 roku.



Rysunek 6. Fragment siatki demograficznej (dane umowne)

Rozdział 2

WSPÓŁCZYNNIKI DEMOGRAFICZNE

2.1. Wprowadzenie

W tym rozdziale przedstawione zostaną zasady budowy współczynników demograficznych definiowanych jako ilorazy zmiennych strumienia i zasobu. Tak rozumiane współczynniki interpretujemy jako liczby zdarzeń demograficznych przypadających na 1 osoborok czasu ekspozycji lub na 1 osoborok życia jednostek zbiorowości.

W celu ułatwienia zapisu i interpretacji, współczynniki demograficzne podaje się niekiedy w przeliczeniu na 1000, 10 000 lub 100 000 osobolat. Można je wówczas interpretować jako liczby zdarzeń przypadających w skali roku na 1000, 10 000 lub 100 000 ludności.

W szerszym znaczeniu współczynniki demograficzne mogą być też konstruowane jako ilorazy dwóch zmiennych zasobu (np. iloraz liczby kobiet do liczby mężczyzn) lub jako ilorazy dwóch zmiennych strumienia (np. iloraz liczby urodzeń do liczby zgonów). Miary tego rodzaju będą nazywane dla odróżnienia wskaźnikami demograficznymi.

2.2. Podział współczynników demograficznych

Metoda wyodrębnienia liczb dożywających oraz zbiorów zdarzeń demograficznych rzutuje na sposób budowy współczynników demograficznych i ich podział. Ze względu na konfigurację danych, wyróżniamy trzy podstawowe ujęcia analizy demograficznej:

- analiza kohortowa lub zamiennie analiza wzdłużna (*Cohort Analysis, Cohort-Age Analysis*),
- analiza przekrojowo-kohortowa (*Period-Cohort Analysis*),
- analiza przekrojowa lub zamiennie analiza transwersalna (*Period Analysis, Period-Age Analysis*).

Analizę kohortową przeprowadza się w odniesieniu do zadanej kohorty lub generacji jednostek, niekiedy dodatkowo w podziale na roczne lub kilkuletnie przedziały czasu własnego, np. wieku. Bierze się tu pod uwagę liczby dożywających I rodzaju i liczby zdarzeń I rodzaju.

Analiza przekrojowo-kohortowa odnosi się do kohorty lub generacji jednostek, w podziale wg ustalonych okresów czasu kalendarzowego, np. lat. W analizie tego rodzaju uwzględnia się liczby dożywających II rodzaju i liczby zdarzeń II rodzaju.

Analiza przekrojowa dotyczy ustalonej zbiorowości, na którą składa się wiele generacji żyjących w tym samym czasie kalendarzowym. Uwzględnia się niekiedy także podział na roczne lub kilkuletnie przedziały czasu własnego, np. wieku. W analizie tej bierze się pod uwagę liczby dożywających II rodzaju i liczby zdarzeń III rodzaju.

Do innych kryteriów klasyfikacji współczynników demograficznych zaliczamy:

- 1) szczegółowość ujęcia badanego zjawiska,
- 2) ryzyko (lub szansę) narażenia jednostek na wystąpienia danego zdarzenia demograficznego,
- 3) oddziaływanie czynników zakłócających.

Ad. 1. Ze względu na szczegółowość ujęcia badanego zjawiska, wyróżniamy:

- **współczynniki cząstkowe (specyficzne, grupowe)** obliczane dla podzbiorowości jednostek, wyodrębnionych z większej zbiorowości ze względu na określone charakterystyki, np. wiek, płeć, miejsce zamieszkania, inne cechy społeczno-ekonomiczne,
- **współczynniki ogólne (surowe)**: obliczane dla całej zbiorowości.

Współczynniki surowe oznaczać będziemy dalej dużymi literami, a współczynniki cząstkowe literami małymi.

Ad. 2. Ze względu na ryzyko lub szansę wystąpienia danego zdarzenia, wyróżniamy:

- **współczynniki pierwszej kategorii**: dotyczą zbiorowości jednostek, które są narażone na ryzyko określonego zdarzenia, np. współczynniki płodności czy zgonów (pojęcie współczynnika demograficznego jest w tym sensie zbieżne z pojęciem „stopy” używanym w analizach finansowych),

- **współczynniki drugiej kategorii:** dotyczą zbiorowości, w której nie wszystkie jednostki są bezpośrednio narażone na ryzyko danego zdarzenia, czego przykładem jest ogólny współczynnik urodzeń.

Ad. 3. Ze względu na wpływ czynników zakłócających, rozważamy:

- **współczynniki netto:** konstruowane z uwzględnieniem wpływu czynników zakłócających,
- **współczynniki brutto:** definiowane z pominięciem czynników zakłócających.

Czynniki zakłócające oddziałują zarówno na liczbę obserwowanych zdarzeń, jak i na czas ekspozycji jednostek zbiorowości. Przykładowo, hipotetyczna liczba urodzeń żywych w ustalonej kohorcie kobiet przy założeniu braku umieralności i braku migracji może różnić się od rzeczywistej liczby urodzeń. Zdarzeniami zakłócającymi są w tym przypadku zgony i migracje kobiet w wieku rozrodczym, powodujące potencjalne „straty” w łącznym czasie ekspozycji, a tym samym w liczbie urodzeń.

Niekiedy zakłada się nioselektywny charakter zdarzeń zakłócających, choć takie założenie nie zawsze jest słuszne. Innymi słowy, zakłada się, że gdyby zdarzenia niezrealizowane w danej zbiorowości były znane, to ich intensywność byłaby taka sama, jak dla zdarzeń zrealizowanych. Przy tym założeniu współczynniki demograficzne obliczone z uwzględnieniem lub bez uwzględniania „utraconych” zdarzeń są jednakowe [por. przykład 3 w sekcji 2.7].

2.3. Współczynniki kohortowe

A. Ogólny, kohortowy współczynnik demograficzny

Rozważmy iloraz liczby zdarzeń demograficznych w pewnej kohorcie s do liczby osobolet życia (lub do łącznego czasu ekspozycji na ryzyko wystąpienia badanego zdarzenia) wśród jednostek tej kohorty. Iloraz taki nazywamy ogólnym, kohortowym współczynnikiem demograficznym. Zapisujemy go ogólnie wzorem

$$W^{koh.s} = \frac{Z^{koh.s}}{K^{koh.s}} \cdot C, \quad (2.1)$$

gdzie:

$Z^{koh.s}$ – liczba zdarzeń demograficznych w s -tej kohorcie,

$K^{koh.s}$ – liczba osobolat życia (łączny czas ekspozycji) jednostek s -tej kohorty,

C – zadana stała.

Iloraz w formule (2.1), podobnie jak wszystkie następne, mnoży się często przez pewną stałą C , np. przez 1000, 10 000, szczególnie wtedy, gdy sam iloraz jest bardzo małym ułamkiem. Zabieg ten ułatwia zapis dziesiętny ułamka, a także interpretację współczynnika [por. przykłady w sekcji 2.7]. Jeśli mnożenie przez stałą nie jest w danym przypadku konieczne, przyjmujemy domyślnie $C = 1$.

B. Cząstkowy (grupowy), kohortowy współczynnik demograficzny

Założmy, że badaną kohortę jednostek podzielono na podzbiorowości z punktu widzenia wybranej cechy społeczno-ekonomicznej. Może to być np. podział ze względu na wiek, płeć, miejsce zamieszkania, wykształcenie, grupy zawodowe itp. W praktyce stosuje się głównie podział ze względu na ustalone grupy wieku $[x, x + n)$ lat.

Przyjmijmy dalej, że x jest indeksem identyfikującym wyodrębnione podzbiorowości.

Kohortowy współczynnik demograficzny dla x -tej podzbiorowości obliczamy jako iloraz liczby zdarzeń demograficznych do liczby osobolat życia (lub do łącznego czasu ekspozycji na ryzyko wystąpienia danego zdarzenia) wśród jednostek tej podzbiorowości. Współczynnik cząstkowy zapiszemy wzorem

$$W_x^{koh.s} = \frac{Z_x^{koh.s}}{K_x^{koh.s}} \cdot C, \quad (2.2)$$

gdzie:

$Z_x^{koh.s}$ – liczba zdarzeń demograficznych, które wystąpiły w x -tej podzbiorowości wyodrębnionej w s -tej kohorcie,

$K_x^{koh.s}$ – liczba osobolat życia (łączny czas ekspozycji) x -tej podzbiorowości w kohorcie s ,

C – zadana stała.

Uwaga 1: Wyjątek w zakresie oznaczeń stanowić będzie współczynnik kohortowy dla podzbiorowości jednostek wyodrębnionych ze względu na przedziały czasu własnego $[x, x+n)$, gdy $n > 1$. Oznaczać będziemy go ogólnie symbolem ${}_nW_x^{koh.s}$. W podobny sposób oznaczone będą wtedy składniki ilorazu po prawej stronie wzoru (2.2), tj. jako ${}_nZ_x^{koh.s}$ oraz ${}_nK_x^{koh.s}$.

Uwaga 2: Współczynniki kohortowe (cząstkowe i ogólne) różnić się mogą pomiędzy kohortami. Poszczególne generacje mogą być umiejscowione w innych okresach historycznych, a z tym wiąże się odmienne środowisko ich życia, np. społeczne, kulturowe czy materialne. Mówimy wówczas o wpływie **efektu kohortowego** na różnice w poziomie współczynników kohortowych.

Uwaga 3: W przeciwieństwie do licznika formuł definiujących współczynnik demograficzny określonego typu, mianownik (liczba osobolat życia lub czas ekspozycji na ryzyko wystąpienia określonego zdarzenia) w praktyce rzadko bywa znany i konieczne staje się zastosowanie pewnych jego przybliżeń.

Załóżmy zatem, że dana zbiorowość została podzielona ze względu na pewne przedziały czasu własnego. Jeśli przyjmiemy, że liczebność w poszczególnych przedziałach zmienia się liniowo wraz z czasem, wówczas w celu aproksymacji liczby osobolat życia lub łącznego czasu ekspozycji w danym przedziale możemy wykorzystać średnią z dwóch liczebności, tj. z początku i końca przedziału lub liczebność zbiorowości w momencie określającym środek przedziału. Liczebność średnią lub środkową mnożymy następnie przez szerokość przedziału wyrażoną w latach.

Przykładowo, rozważmy cząstkowy współczynnik kohortowy (2.2), w którym indeks dolny x oznacza roczne przedziały wieku $[x, x+1)$ lat. Wówczas formuła aproksymująca mianownik $K_x^{koh.s}$ może być zapisana jako średnia arytmetyczna z liczb dożywających I rodzaju postaci

$$K_x^{koh.s} \approx \frac{1}{2} \cdot (U_x^{koh.s} + U_{x+1}^{koh.s}) \quad (2.3)$$

lub też jako liczba $U_{x+\frac{1}{2}}^{koh.s}$ dożywających do połowy przedziału $[x, x+1)$, czyli osiagających wiek $x + \frac{1}{2}$ lat

$$K_x^{(koh.s)} \approx U_{x+\frac{1}{2}}^{(koh.s)}. \quad (2.4)$$

Wtedy współczynnik kohortowy $W_x^{(koh.s)}$ przybliżamy wzorami

$$W_x^{koh.s} \approx \frac{Z_x^{koh.s}}{\frac{1}{2} \cdot (U_x^{koh.s} + U_{x+1}^{koh.s})} \cdot C \quad (2.5)$$

lub

$$W_x^{koh.s} \approx \frac{Z_x^{koh.s}}{U_{x+\frac{1}{2}}^{koh.s}} \cdot C. \quad (2.6)$$

W przypadku przedziałów wieku postaci $[x, x + n)$, gdy $n > 1$ ($n \in \mathbf{N}$), analogiczne przybliżenia liczby osobolat życia ${}_nK_x^{(koh.s)}$ mają postać

$${}_nK_x^{(koh.s)} \approx n \cdot \frac{U_x^{koh.s} + U_{x+n}^{koh.s}}{2}, \quad (2.7)$$

lub

$${}_nK_x^{(koh.s)} \approx n \cdot U_{x+\frac{n}{2}}^{(koh.s)}, \quad (2.8)$$

gdzie n oznacza szerokość przedziału.

Współczynnik kohortowy ${}_nW_x^{(koh.s)}$ aproksymujemy wtedy na podstawie wzorów

$${}_nW_x^{koh.s} \approx \frac{nZ_x^{koh.s}}{\frac{n}{2} \cdot (U_x^{koh.s} + U_{x+n}^{koh.s})} \cdot C \quad (2.9)$$

lub

$${}_nW_x^{koh.s} \approx \frac{nZ_x^{koh.s}}{n \cdot U_{x+\frac{n}{2}}^{koh.s}} \cdot C. \quad (2.10)$$

2.4. Współczynniki przekrojowo-kohortowe

A. Ogólny, przekrojowo-kohortowy współczynnik demograficzny

Rozważmy iloraz liczby zdarzeń demograficznych, które wystąpiły w kohorcie s w przedziale czasu kalendarzowego t do liczby osobolat życia (lub do łącznego czasu ekspozycji na ryzyko wystąpienia badanego zdarzenia) wśród jednostek tej kohorty w okresie t . Iloraz ten

nazywamy ogólnym, przekrojowo-kohortowym współczynnikiem demograficznym. Zapisujemy go wzorem

$$W^{koh.s}(t) = \frac{Z^{koh.s}(t)}{K^{koh.s}(t)} \cdot C, \quad (2.11)$$

gdzie:

$Z^{koh.s}(t)$ – liczba zdarzeń demograficznych w przedziale czasu kalendarzowego t w kohorcie s ,

$K^{koh.s}(t)$ – liczba osobolat życia (łączy czas ekspozycji) w okresie t dla jednostek s -tej kohorty,

C – zadana stała.

B. Częstkowy (grupowy), przekrojowo-kohortowy współczynnik demograficzny

Podobnie jak w przypadku grupowych współczynników kohortowych, możemy rozważać przekrojowo-kohortowe współczynniki cząstkowe.

Załóżmy, że daną kohortę jednostek podzielono z punktu widzenia wybranej cechy społeczno-ekonomicznej na pewne podzbiorowości. Niech x będzie indeksem identyfikującym te podzbiorowości. Przekrojowo-kohortowy współczynnik demograficzny dla x -tej podzbiorowości zapiszemy wzorem

$$W_x^{koh.s}(t) = \frac{Z_x^{koh.s}(t)}{K_x^{koh.s}(t)} \cdot C, \quad (2.12)$$

gdzie:

$Z_x^{koh.s}(t)$ – liczba zdarzeń demograficznych, które wystąpiły w okresie t w x -tej podzbiorowości wyodrębnionej w s -tej kohorcie,

$K_x^{koh.s}(t)$ – liczba osobolat życia w okresie t dla x -tej podzbiorowości w kohorcie s ,

C – zadana stała.

Uwaga 4: W przypadku szczególnym, jeśli podzbiorowości zostały wyodrębnione ze względu na roczne przedziały czasu własnego $[x, x+1)$, a t oznacza ustalony rok kalendarzowy, wówczas (2.12) sprowadza się do (2.11), ponieważ przedział czasu własnego w ustalonej kohorcie s w roku t jest z góry zdeterminowany, stąd indeks x można pominąć.

Uwaga 5: Analiza przekrojowo-kohortowa, podobnie jak analiza kohortowa, umożliwia porównania pomiędzy kohortami, w tym badanie wpływu **efektu kohortowego**. Analiza taka ma dodatkowo tę zaletę, że może być ograniczona do ostatniego okresu kalendarzowego, np. do ostatniego roku lub kilku ostatnich lat, a więc z pominięciem okresów odległych w czasie.

Uwaga 6: Ze względu na ograniczony dostęp do dokładnych danych dotyczących czasu ekspozycji lub liczby osobolat życia zbiorowości w danym przedziale czasu kalendarzowego, mianownik wzorów (2.11) lub (2.12) przybliżamy zwykle za pomocą średniej arytmetycznej z liczb dożywających II rodzaju odnotowanych na początku i końcu analizowanego okresu (liczebność średnia) lub też za pomocą liczby dożywających II rodzaju odnotowanej w połowie danego okresu kalendarzowego (liczebność środkowa). W obu sytuacjach liczebność średnią lub środkową mnożymy przez długość okresu wyrażoną w latach.

Przykładowo, formuła aproksymacji ogólnego współczynnika kohortowo-przekrojowego $W^{koh.s}(t)$ dla zadanego przedziału czasu kalendarzowego $t = [\tau_1, \tau_2]$ o długości $\tau_2 - \tau_1$ jest postaci

$$W^{koh.s}(t) \approx \frac{Z^{koh.s}(t)}{(\tau_2 - \tau_1) \cdot \frac{L^{koh.s}(\tau_1) + L^{koh.s}(\tau_2)}{2}} \cdot C, \quad (2.13)$$

lub

$$W^{koh.s}(t) \approx \frac{Z^{koh.s}(t)}{(\tau_2 - \tau_1) \cdot L^{koh.s}(\bar{\tau})} \cdot C, \quad (2.14)$$

gdzie $L^{koh.s}(\tau_1)$, $L^{koh.s}(\tau_2)$, $L^{koh.s}(\bar{\tau})$ są liczebnościami kohorty odpowiednio na początku, końcu i w środku okresu t . Gdy t jest ustalonym rokiem kalendarzowym, wówczas τ_1 , τ_2 oznaczają początek i koniec roku t , $\bar{\tau}$ jest środkiem roku, natomiast $\tau_2 - \tau_1 = 1$.

Przez analogię do (2.13), (2.14), cząstkowy współczynnik kohortowo-przekrojowy dla x -tej podzbiorowości przybliżamy za pomocą wzorów

$$W_x^{koh.s}(t) \approx \frac{Z_x^{koh.s}(t)}{(\tau_2 - \tau_1) \cdot \frac{L_x^{koh.s}(\tau_1) + L_x^{koh.s}(\tau_2)}{2}} \cdot C, \quad (2.15)$$

lub

$$W_x^{koh.s}(t) \approx \frac{Z_x^{koh.s}(t)}{(\tau_2 - \tau_1) \cdot L_x^{koh.s}(\bar{\tau})} \cdot C. \quad (2.16)$$

W szczególnym przypadku indeks podzbiorowości x może być oznaczeniem przedziałów wieku.

2.5. Współczynniki przekrojowe

A. Ogólny, przekrojowy współczynnik demograficzny

Rozważmy iloraz liczby określonych zdarzeń demograficznych, które wystąpiły w zbiorowości żyjących w danym przedziale czasu kalendarzowego t do liczby osobolat życia (lub do łącznego czasu ekspozycji na ryzyko wystąpienia badanego zdarzenia) w zbiorowości ogółu żyjących w okresie t . Iloraz taki nazywamy ogólnym, przekrojowym współczynnikiem demograficznym. Zapisujemy go wzorem

$$W(t) = \frac{Z(t)}{K(t)} \cdot C, \quad (2.17)$$

gdzie:

$Z(t)$ – liczba zdarzeń odnotowanych w okresie kalendarzowym t ,

$K(t)$ – liczba osobolat życia w okresie t ogółu żyjących w tym okresie,

C – zadana stała.

B. Częstkowy (grupowy), przekrojowy współczynnik demograficzny

Żałómy, że zbiorowość żyjących w okresie t podzielono na pewne podzbiorowości z punktu widzenia wybranej cechy społeczno-ekonomicznej. Niech x będzie indeksem identyfikującym wyodrębnione podzbiorowości. Przekrojowy, cząstkowy współczynnik demograficzny dla x -tej podzbiorowości w okresie t dany jest wzorem

$$W_x(t) = \frac{Z_x(t)}{K_x(t)} \cdot C, \quad (2.18)$$

gdzie:

$Z_x(t)$ – liczba zdarzeń demograficznych, które wystąpiły w x -tej podzbiorowości w okresie t ,

$K_x(t)$ – liczba osobolat życia (łączny czas ekspozycji) w okresie t jednostek x -tej podzbiorowości,

C – zadana stała.

Uwaga 7: Analiza przekrojowa umożliwia badanie zdarzeń demograficznych w zbiorowości jednostek żyjących w tym samym czasie i porównywanie rezultatów w wybranych odstępach czasowych. Pozwala to uchwycić wpływ **efektu kalendarzowego** na różnice w poziomach przekrojowych współczynników demograficznych.

Uwaga 8: W praktyce pojawia się zwykle konieczność przybliżania mianownika formuł definiujących współczynniki demograficzne w zadanym przedziale czasu kalendarzowego [zob. też Uwagi 3 i 6].

Założmy, że liczebność zbiorowości w danym przedziale czasu kalendarzowego zmienia się liniowo. Wówczas mianownik współczynnika (2.17) dla okresu $t = [\tau_1, \tau_2]$ o długości $\tau_2 - \tau_1$ przybliżamy wzorami

$$K(t) \approx (\tau_2 - \tau_1) \cdot \frac{L(\tau_1) + L(\tau_2)}{2} \quad (2.19)$$

lub

$$K(t) \approx (\tau_2 - \tau_1) \cdot L(\bar{\tau}). \quad (2.20)$$

Stąd przybliżona formuła na współczynnik (2.17) przybiera postać

$$W(t) \approx \frac{Z(t)}{(\tau_2 - \tau_1) \cdot \frac{L(\tau_1) + L(\tau_2)}{2}} \cdot C \quad (2.21)$$

lub

$$W(t) \approx \frac{Z(t)}{(\tau_2 - \tau_1) \cdot L(\bar{\tau})} \cdot C, \quad (2.22)$$

gdzie $L(\tau_1)$, $L(\tau_2)$, $L(\bar{\tau})$ są liczebnościami zbiorowości odpowiednio na początku, końcu i w środku okresu t . Gdy t jest zadanym rokiem, wtedy τ_1 , τ_2 , $\bar{\tau}$ oznaczają początek, koniec i środek roku, natomiast $\tau_2 - \tau_1 = 1$.

Analogicznie, cząstkowy współczynnik przekrojowy (2.18) dla x -tej podzbiorowości w okresie $t = [\tau_1, \tau_2]$ przybliżamy za pomocą formuł

$$W_x(t) \approx \frac{Z_x(t)}{(\tau_2 - \tau_1) \cdot \frac{L_x(\tau_1) + L_x(\tau_2)}{2}} \cdot C \quad (2.23)$$

lub

$$W_x(t) \approx \frac{Z_x(t)}{(\tau_2 - \tau_1) \cdot L_x(\bar{\tau})} \cdot C. \quad (2.24)$$

W szczególnym przypadku indeks x może oznaczać ustalony przedział wieku.

Uwaga 9: Inna metoda określania liczby osobolat życia lub łącznego czasu ekspozycji w zadanym okresie odwołuje się do pojęcia stopy wzrostu populacji [zob. rozdział 6 – formuła (6.34)].

2.6. Wybrane współczynniki demograficzne

W rozważaniach tej sekcji dotyczących wybranych współczynników demograficznych przyjmujemy, iż stała C jest równa 1. Zakładamy także, że czasem własnym jednostek jest wiek mierzony w latach. Rozważaną zbiorowością będzie ustalona kohorta urodzeniowa s lub też ogół jednostek żyjących w ustalonym okresie kalendarzowym t .

Dotychczasowe oznaczenia zawierały symbol kohorty s i symbol okresu t dla odróżnienia ujęcia kohortowego od ujęcia przekrojowego. Aby uprościć zapis, wspomniane oznaczenia zostaną w tej sekcji pominięte, ponieważ prezentowane dalej współczynniki definiowane będą w podobny sposób w obu ujęciach, tj. można je odnieść do zbiorowości, którą jest kohorta lub ogół żyjących w danym okresie. Zaznaczyć jednak należy, że współczynniki kohortowe i przekrojowe w ogólnym przypadku co do wartości **nie są jednakowe**, poza nielicznymi wyjątkami, o których będzie mowa w rozdziale 3.

2.6.1. Ogólny współczynnik urodzeń (rodności)

Ogólny współczynnik urodzeń CBR (*crude birth rate*) jest ilorazem liczby urodzeń żywych U_0 w danej zbiorowości do łącznej liczby osobolet życia K tej zbiorowości, co zapisujemy wzorem

$$CBR = \frac{U_0}{K}, \quad (2.25)$$

gdzie

$$K = \sum_{x=0}^{X-1} K_x. \quad (2.26)$$

Symbol X w (2.26) określa górną granicę wieku (przyjmujemy dalej $X = 120$ lat).

CBR w podanej formule informuje, ile urodzeń żywych przypada na 1 osoborok życia zbiorowości.

W przypadku kohorty urodzeniowej, współczynnik urodzeń jest ilorazem początkowej liczebności do łącznej liczby osobolet życia kohorty. Częściej jednak CBR jest rozważany w odniesieniu do zbiorowości ogółu żyjących w tym samym okresie kalendarzowym, np. w tym samym roku. Wówczas licznik równy jest liczbie urodzeń żywych w danym roku, a mianownik wyraża łączną liczbą osobolet życia w ciągu roku wszystkich jednostek danej zbiorowości.

Ponieważ na mianownik współczynnika CBR składa się liczba osobolat życia zarówno mężczyzn, jak i kobiet, dlatego, zgodnie z klasyfikacją podaną w sekcji 2.2, współczynnik urodzeń zaliczamy do współczynników drugiej kategorii.

2.6.2. Ogólny współczynnik płodności

Ogólny współczynnik płodności GFR (*general fertility rate*) jest ilorazem liczby żywo urodzonych dzieci B do łącznego czasu ekspozycji $K^{(fem)}$ kobiet w wieku rozrodczym na szansę urodzenia dziecka, co zapisujemy wzorem

$$GFR = \frac{B}{K^{(fem)}}, \quad (2.27)$$

gdzie

$$K^{(fem)} = \sum_{x=a_1}^{a_2-1} K_x^{(fem)}. \quad (2.28)$$

Liczby a_1 , a_2 w (2.28) wyznaczają przedział $[a_1, a_2)$ wieku rozrodczego (przyjmujemy $a_1 = 15$, $a_2 = 50$ lat).

Współczynnik (2.27) informuje, ile żywo urodzonych dzieci przypada na 1 osoborok życia kobiety w wieku rozrodczym. Może być on konstruowany dla ustalonej kohorty kobiet lub dla kobiet z różnych generacji żyjących w tym samym okresie kalendarzowym. W drugim przypadku $B = U_0$.

Współczynnik GFR zalicza się do mierników pierwszej kategorii, bowiem odnosi badane zdarzenie (urodzenie dziecka) do populacji kobiet, podlegających „ryzyku” (szansie) jego wystąpienia.

2.6.3. Ogólny współczynnik zgonów

Ogólny współczynnik zgonów CDR (*crude death rate*) jest ilorazem liczby zgonów D do łącznego czasu ekspozycji K danej zbiorowości na ryzyko zgonu, czyli

$$CDR = \frac{D}{K}, \quad (2.29)$$

gdzie

$$K = \sum_{x=0}^{X-1} K_x, \quad (2.30)$$

przy czym X oznacza górną granicę wieku.

Iloraz (2.29) informuje o liczbie zgonów przypadających na 1 osoborok życia zbiorowości.

W przypadku kohorty urodzeniowej zachodzi $CBR = CDR$, ponieważ liczebność początkowa kohorty równa jest łącznej liczbie zgonów. W praktyce zwykle stosowane jest ujęcie przekrojowe, tj. liczbę zgonów D ustala się dla danego roku kalendarzowego i odnosi do czasu ekspozycji K na ryzyko zgonu w ciągu roku dla ogółu jednostek żyjących w tym samym czasie.

2.6.4. Współczynnik przyrostu naturalnego

Współczynnik przyrostu naturalnego RNI (*rate of natural increase*) jest różnicą pomiędzy ogólnymi (surowymi) współczynnikami urodzeń i zgonów

$$RNI = CBR - CDR. \quad (2.31)$$

Współczynnik RNI interpretujemy jako wielkość przyrostu (ubytku) zbiorowości spowodowanego ruchem naturalnym w przeliczeniu na 1 osoborok życia jednostek zbiorowości. Dodatnie wartości RNI wskazują na nadwyżkę liczby urodzeń nad liczbą zgonów, a wartości ujemne – na relację odwrotną.

Warto zauważyć, że w przypadku kohorty urodzeniowej $RNI = 0$. Z tego względu współczynnik przyrostu naturalnego obliczamy głównie na podstawie danych przekrojowych.

2.6.5. Cząstkowe współczynniki płodności

Cząstkowy współczynnik płodności (*age-specific fertility rate*) dla danej grupy wieku $[x, x + n)$ oznaczać będziemy symbolem ${}_n f_x$. Definiujemy go jako iloraz liczby ${}_n B_x$ żywo urodzonych dzieci w zbiorowości kobiet będących w wieku $[x, x + n)$ lat do łącznego czasu ekspozycji ${}_n K_x^{(fem)}$ kobiet na szansę urodzenia dziecka w tej grupie wieku, czyli

$${}_n f_x = \frac{{}_n B_x}{{}_n K_x^{(fem)}}. \quad (2.32)$$

Dla $n = 1$ zapis można uprościć do postaci

$$f_x = \frac{B_x}{K_x^{(fem)}}. \quad (2.33)$$

Cząstkowe współczynniki płodności mogą być wyznaczane zarówno w ujęciu kohortowym, jak i przekrojowym. W pierwszym przypadku

zbiorowość stanowi kohorta kobiet, w drugim – kobiety z różnych generacji żyjące w tym samym czasie. Zakładamy przy tym, że analizowana grupa wieku $[x, x + n)$ zawiera się w przedziale wieku prokreacyjnego $[15, 50)$ lat. W praktyce rozważa się na ogół podział na grupy roczne lub pięcioletnie.

W przypadku cząstkowych współczynników płodności analizujemy wyłącznie zbiorowość kobiet podlegających szansie wystąpienia danego zdarzenia. Zgodnie z klasyfikacją podaną w sekcji 2.2, cząstkowe współczynniki płodności zaliczamy do grupy mierników pierwszej kategorii.

2.6.6. Cząstkowe współczynniki zgonów

Cząstkowy współczynnik zgonów (*age-specific mortality rate*) dla grupy wieku $[x, x + n)$ lat oznaczamy symbolem ${}_n m_x$. Definiujemy go jako iloraz liczby zgonów ${}_n D_x$ odnotowanych w zbiorowości osób w wieku $[x, x + n)$ lat do łącznego czasu ekspozycji ${}_n K_x$ na ryzyko zgonu w tej grupie wieku, czyli

$${}_n m_x = \frac{{}_n D_x}{{}_n K_x}. \quad (2.34)$$

Dla $n = 1$ zapis można uprościć do postaci

$$m_x = \frac{D_x}{K_x}. \quad (2.35)$$

Podobnie jak w przypadku współczynników płodności, populację w ujęciu kohortowym stanowi ustalona kohorta osób, a w ujęciu przekrojowym – osoby w wieku $[x, x + n)$ lat żyjące jednocześnie w ustalonym czasie.

Uwaga 10: Szczególny rodzaj współczynnika przekrojowego stanowi współczynnik zgonów niemowląt, mierzący natężenie zgonów w grupie dzieci w wieku $[0, 1)$ lat lub, innymi słowy, w wieku 0 ukończonych lat. Oznaczmy ten współczynnik symbolem $m_0(t)$. Jedną z formuł stosowaną do przybliżania $m_0(t)$ ma postać

$$m_0(t) \approx \frac{D_0(t)}{\alpha U_0(t) + (1 - \alpha) U_0(t - 1)} \cdot C, \quad (2.36)$$

gdzie $D_0(t)$ jest liczbą zgonów niemowląt w roku t , $U_0(t)$ oznacza urodzenia żywe w roku t , α to wskaźnik udziału zgonów niemowląt urodzonych i zmarłych w roku t w ogólnej liczbie niemowląt zmarłych w tym roku, natomiast C jest zadaną stałą.

Często stosuje się wzór, w którym $\alpha = 1$. Mamy wówczas

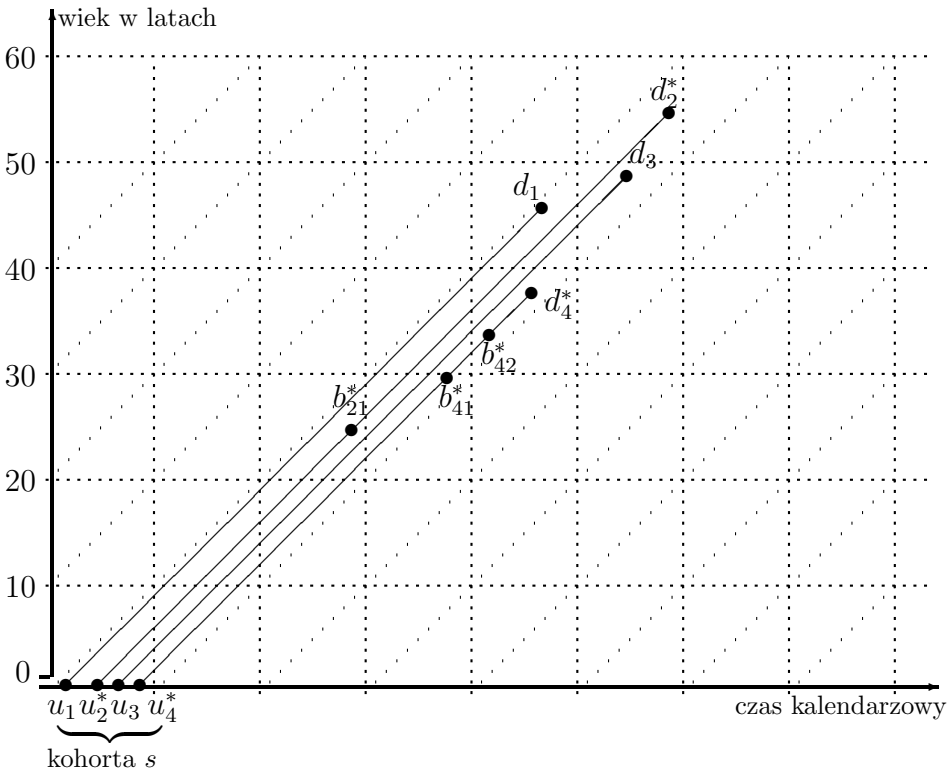
$$m_0(t) \approx \frac{D_0(t)}{U_0(t)} \cdot C. \tag{2.37}$$

2.7. Przykłady

Przykład 3. Rozważmy umowną kohortę urodzeniową składającą się z dwóch kobiet i dwóch mężczyzn, których linie życia przedstawia rysunek 7.

Niech punkt oznaczony symbolem u_k na początku k -tej linii oznacza urodzenie się k -tej osoby w kohorcie, d_k niech oznacza zgon k -tej osoby, natomiast b_{ki} – urodzenie i -tego dziecka przez k -tą osobę (kobietę). Zdarzenia dotyczące kobiet oznaczono dodatkowo gwiazdką $*$.

Obliczymy ogólny współczynnik zgonów CDR , urodzeń CBR oraz współczynnik płodności GFR .



Rysunek 7. Fragment siatki demograficznej i przykładowe linie życia

Na podstawie rysunku 7 możemy łatwo ustalić, że początkowa liczebność U_0 kohorty oraz liczba zgonów D są równe i wynoszą 4, natomiast liczba urodzonych dzieci B w tej kohorcie wynosi 3. Z kolei łączny czas ekspozycji na ryzyko zgonu jest równy

$$K = 46 + 55 + 49 + 38 = 188 \text{ osobolat,}$$

a łączny czas ekspozycji kobiet w wieku rozrodczym na szansę urodzenia dziecka wynosi

$$K^{(fem)} = 35 + 23 = 58 \text{ osobolat.}$$

Wartości składające się na liczbę K oznaczają lata życia czterech osób należących do rozważanej kohorty, natomiast w przypadku $K^{(fem)}$ – czas ekspozycji dwóch kobiet mierzony od momentu wejścia w wiek prokreacyjny, tj. przedział wieku [15, 50) lat, do momentu zakończenia okresu rozrodczego (zakładamy dodatkowo ciągłość czasu ekspozycji w tym okresie) lub do momentu zgonu, o ile zdarzenie to nastąpiło wcześniej.

Na tej podstawie otrzymujemy ogólny współczynnik zgonów i jednocześnie ogólny współczynnik urodzeń (rodności)

$$CDR = CBR = \frac{4}{188},$$

a także ogólny współczynnik płodności

$$GFR = \frac{3}{58}.$$

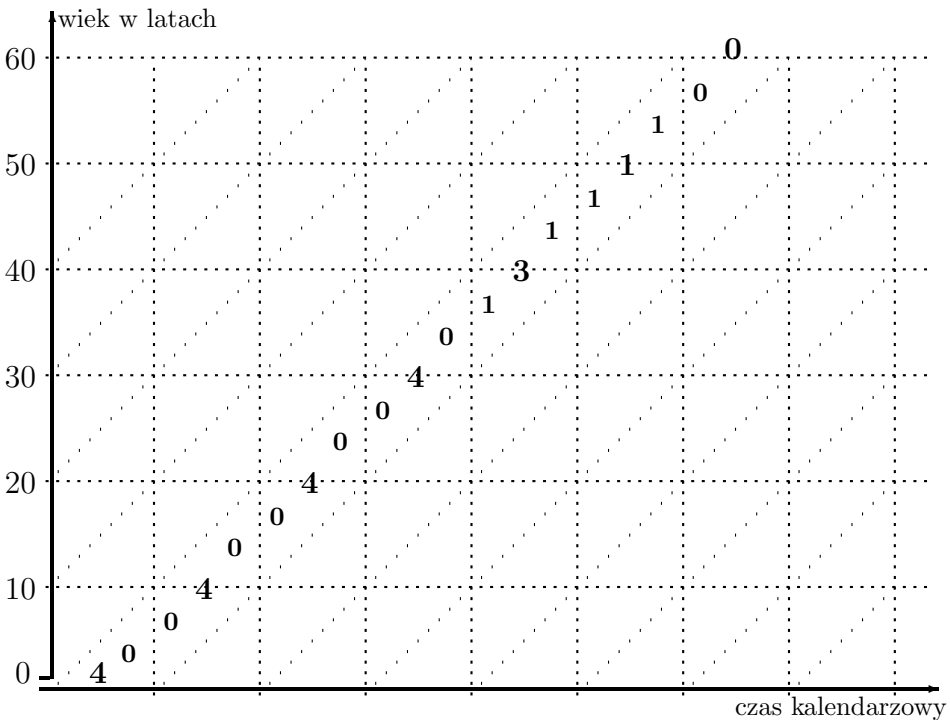
Warto zauważyć, że w obliczeniach czasu ekspozycji $K^{(fem)}$ dla obu kobiet uwzględnione zostało wystąpienie czynnika zakłócającego, jakim jest zgon jednej z nich, który skrócił o 12 lat czas ekspozycji tej kobiety na szansę urodzenia dziecka. Jeśli przyjmiemy, że natężenie płodności w relacji do tej części czasu wynosi $3/58$, to współczynnik GFR wyznaczony z uwzględnieniem „utraconego” czasu ekspozycji osiągnie tę samą wartość. Mamy bowiem

$$GFR = \frac{3 + 12 \cdot \frac{3}{58}}{58 + 12} = \frac{3(58 + 12)}{58(58 + 12)} = \frac{3}{58},$$

przy czym składnik $12 \cdot \frac{3}{58}$ jest oszacowaniem dodatkowej liczby urodzeń przypadającej na „utracony” czas ekspozycji.

Do wyznaczenia rozważanych wyżej współczynników korzystaliśmy z danych dotyczących m.in. czasu ekspozycji. Jeśli zamiast tego posiadamy informacje o liczbie dożywających I rodzaju oraz o liczbie zdarzeń I rodzaju (urodzeń lub zgonów), wtedy możemy wyznaczyć odpowiednie przybliżenia.

Rysunek 8 zawiera zagregowane informacje o liczbach dożywających I rodzaju (na odcinkach poziomych siatki) i liczbach zgonów (wewnątrz trójkątnych pól) ustalone na podstawie rysunku 7.



Rysunek 8. Fragment siatki demograficznej, liczby dożywających I rodzaju oraz liczby zgonów (dane na podstawie rysunku 7)

Przybliżymy łączny czas ekspozycji K na ryzyko zgonu dla wszystkich jednostek kohorty. Mamy

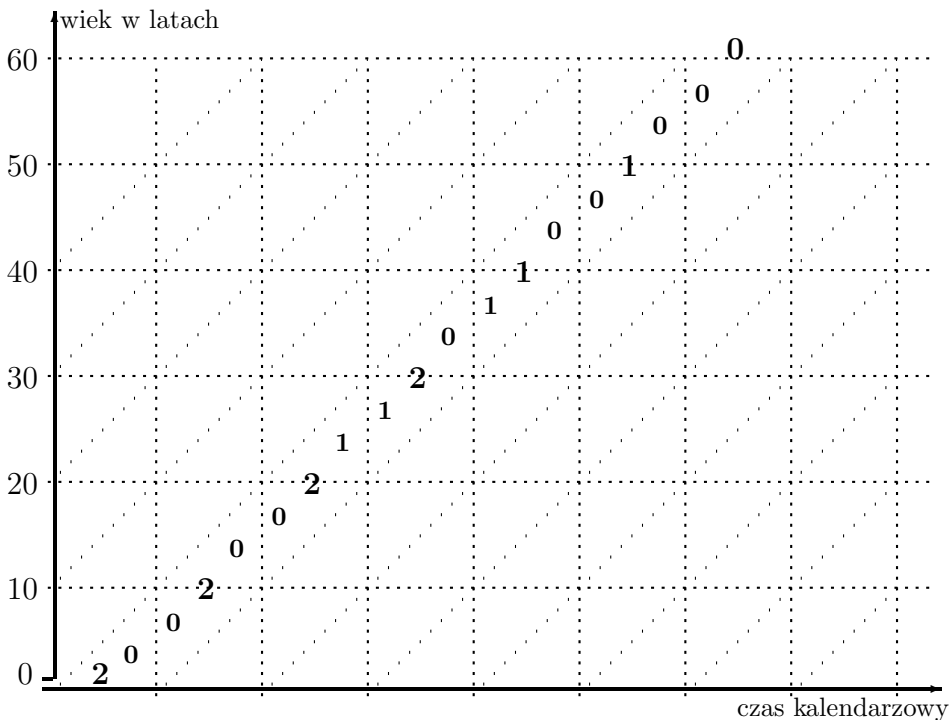
$$K \approx \left[4 + 4 + 4 + \frac{4+3}{2} + \frac{3+1}{2} + \frac{1+0}{2} \right] \cdot 10 = 180 \text{ osobolat,}$$

gdzie 10 (lat) oznacza jednostkę czasu, tj. długość poszczególnych przedziałów czasu własnego i kalendarzowego.

Stąd otrzymujemy

$$CDR = CBR \approx \frac{4}{180}.$$

Rozważmy teraz zagregowane informacje o liczbach kobiet dożywających określonego wieku (liczby dożywających I rodzaju) i liczbach urodzonych dzieci (liczby wewnątrz trójkątów Lexisa). Układ danych przedstawia rysunek 9.



Rysunek 9. Fragment siatki demograficznej, liczby dożywających I rodzaju oraz liczby urodzeń (dane na podstawie rysunku 7)

Przybliżony łączny czas ekspozycji $K^{(fem)}$ kobiet w wieku rozrodczym na szansę urodzenia dziecka wyznaczmy na podstawie danych z rysunku 9. Mamy

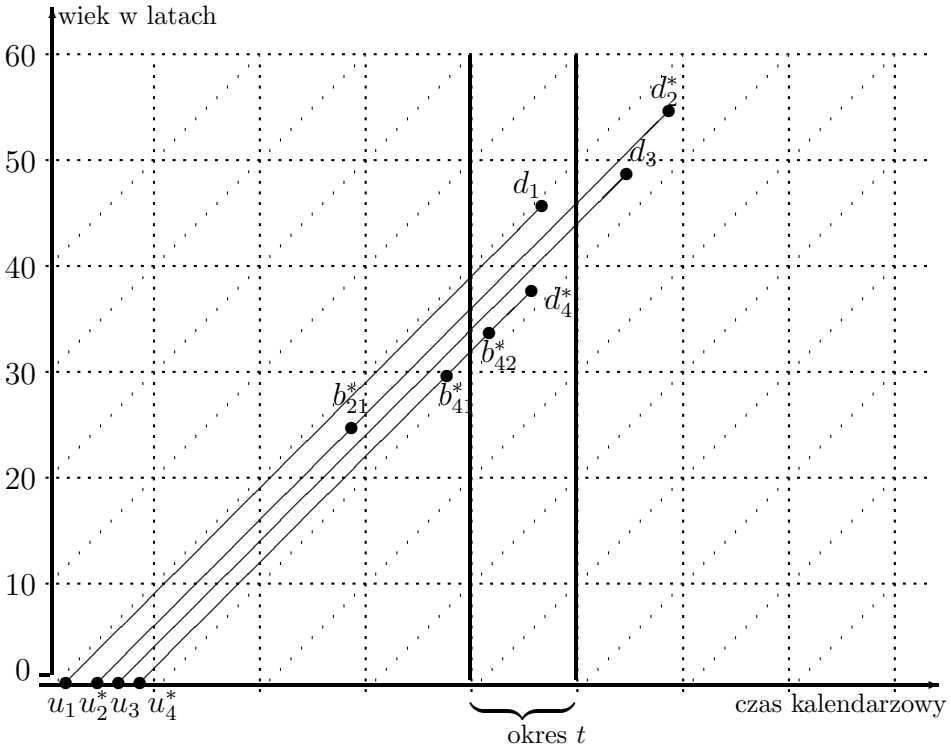
$$K^{(fem)} \approx 2 \cdot 5 + \left[2 + \frac{2+1}{2} + 1 \right] \cdot 10 = 55 \text{ osobolat.}$$

Stąd

$$GFR \approx \frac{3}{55}.$$

Warto zauważyć, że wyniki uzyskane na podstawie danych zagregowanych są przybliżeniem wartości poszczególnych współczynników CDR, CBR, GFR otrzymanych wcześniej z danych indywidualnych.

Przykład 4. Na podstawie danych z przykładu 3 obliczymy ogólne współczynniki zgonów i płodności, ograniczając rozważania do wybranego przedziału czasu kalendarzowego t , zaznaczonego na rysunku 10 [oznaczenia analogiczne, jak w przykładzie 3].



Rysunek 10. Fragment siatki demograficznej, przykładowe linie życia i wybrany przedział czasu kalendarzowego t

Liczba zgonów D i liczba urodzeń B w zadanym okresie kalendarzowym (tutaj 10-letnim) są równe

$$D = 2, \quad B = 1.$$

Z kolei czas ekspozycji na ryzyko zgonu w tym okresie wynosi

$$K \approx 8 + 10 + 10 + 6 = 34 \text{ osobołata,}$$

a czas ekspozycji kobiet w wieku rozrodczym na szansę urodzenia dziecka jest równy

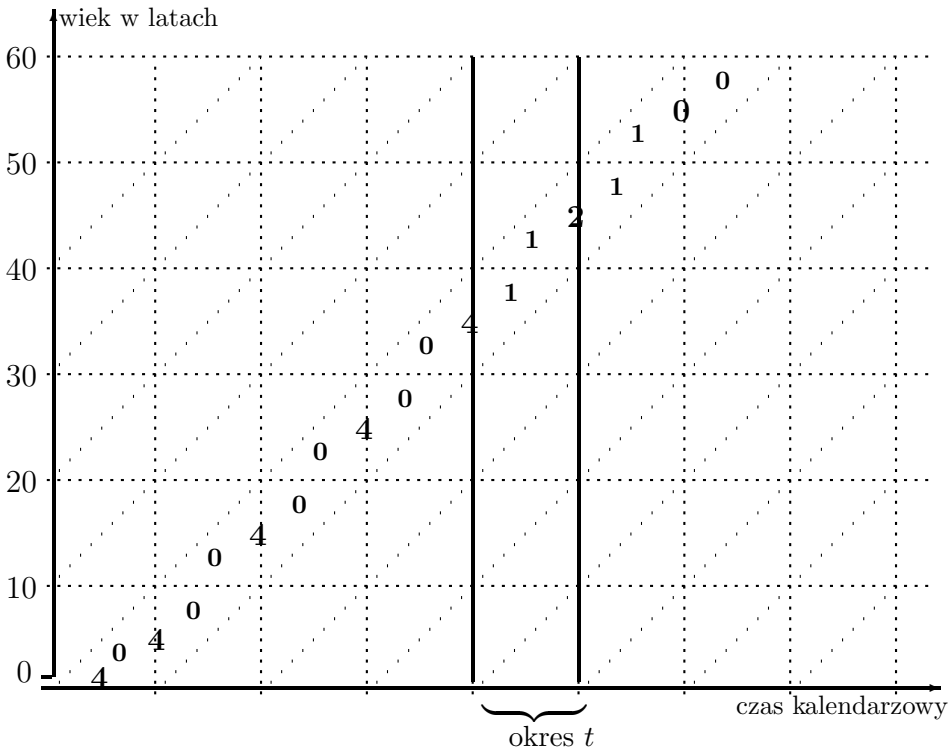
$$K^{(fem)} \approx 10 + 8 = 18 \text{ osobolat,}$$

stąd współczynniki zgonów i płodności dla rozważanego okresu wynoszą odpowiednio

$$CDR = \frac{2}{34},$$

$$GFR = \frac{1}{18}.$$

Jeśli dysponujemy danymi zagregowanymi, tj. posiadamy informacje o liczbach dożywających II rodzaju i liczbach zdarzeń II rodzaju (tu: zgonów lub urodzeń), wówczas możliwe jest wyznaczenie przybliżeń analizowanych współczynników.



Rysunek 11. Fragment siatki demograficznej, liczby dożywających II rodzaju oraz liczby zgonów (dane na podstawie rysunku 10)

Rysunek 11 ilustruje liczby dożywających II rodzaju (na odcinkach pionowych siatki) oraz liczby zgonów w badanej kohorcie (wewnątrz trójkątnych pól).

Liczba zgonów D i czas ekspozycji K na ryzyko zgonu w rozważanym okresie kalendarzowym wynoszą odpowiednio

$$D = 2,$$

$$K \approx \frac{4 + 2}{2} \cdot 10 = 30 \text{ osobolat},$$

zatem współczynnik CDR obliczony dla zadanego okresu równy jest w przybliżeniu

$$CDR \approx \frac{2}{30}.$$

Wynik jest aproksymacją współczynnika CDR wyznaczonego z danych indywidualnych.

Do obliczenia współczynnika płodności GFR skorzystamy z liczb dożywających II rodzaju dla kobiet oraz liczb urodzeń (rysunek 12). Różnice w stosunku do poprzedniej siatki dotyczą określenia zdarzeń demograficznych, którymi są obecnie urodzenia.

Wyznaczenie GFR na podstawie zagregowanych danych rozpoczynamy od ustalenia liczby urodzeń B oraz czasu ekspozycji $K^{(fem)}$ kobiet na szansę urodzenia dziecka w zadanym okresie kalendarzowym. Mamy na podstawie rysunku 12

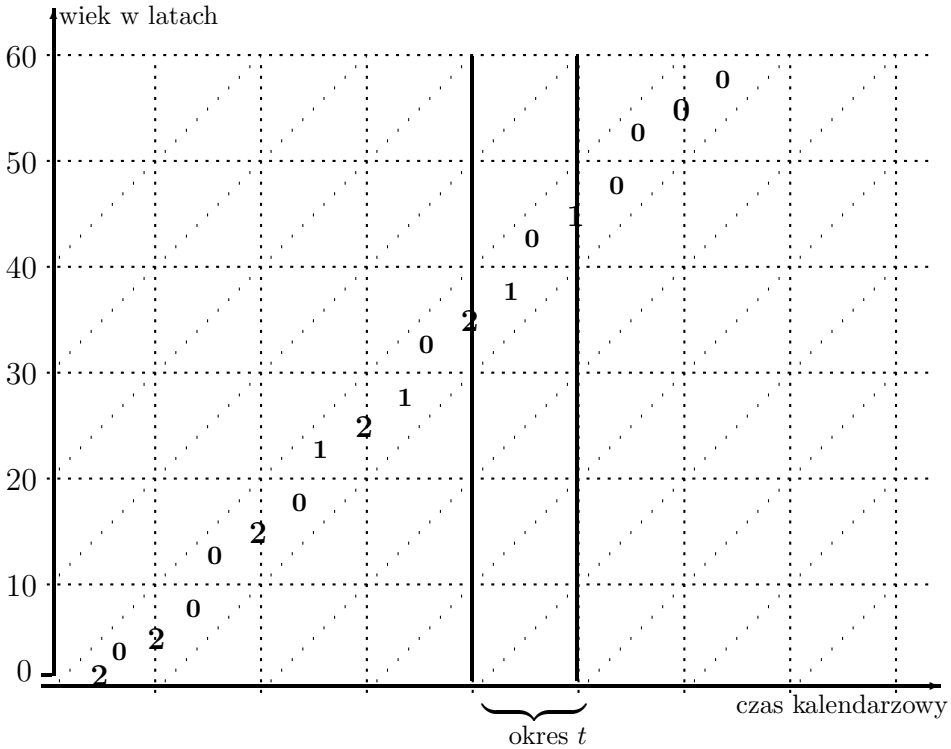
$$B = 1,$$

$$K^{(fem)} \approx \frac{2 + 1}{2} \cdot 10 = 15 \text{ osobolat}.$$

Stąd otrzymujemy

$$GFR \approx \frac{1}{15}.$$

Wynik jest przybliżeniem współczynnika GFR uzyskanego z danych indywidualnych.



Rysunek 12. Fragment siatki demograficznej, liczba dożywających II rodzaju (kobiety) oraz liczby urodzeń (dane na podstawie rysunku 10)

Przykład 5. Dla kohorty z przykładu 3 obliczymy cząstkowy współczynnik płodności i cząstkowy współczynnik zgonów w grupie wieku $[30, 40)$ lat (por. rysunek 13).

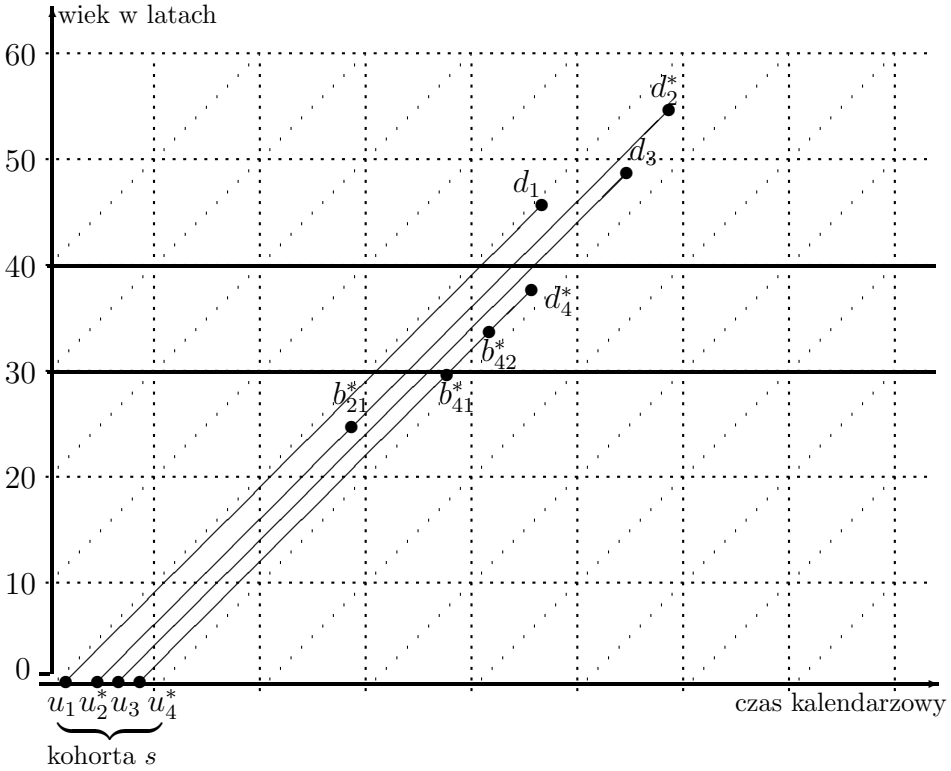
W przypadku cząstkowych współczynników płodności analizę ograniczamy do kobiet. W tej kohorcie są tylko dwie kobiety, obie przeżyły okres odpowiednio 10 lat i 8 lat swojego życia w przedziale wieku $[30, 40)$ lat.

Oznaczmy łączny czas ekspozycji obu kobiet na szansę urodzenia dziecka w zadanym przedziale wieku i liczbę urodzonych w tym czasie dzieci symbolami odpowiednio ${}_{10}K_{30}^{(fem)}$ oraz ${}_{10}B_{30}$. Mamy

$${}_{10}B_{30} = 1, \quad {}_{10}K_{30}^{(fem)} = 10 + 8 = 18 \text{ osobolat.}$$

Zatem analizowany cząstkowy współczynnik płodności wynosi

$${}_{10}f_{30} = \frac{1}{18}.$$



Rysunek 13. Fragment siatki demograficznej, przykładowe linie życia i wybrany przedział wieku [30, 40) lat

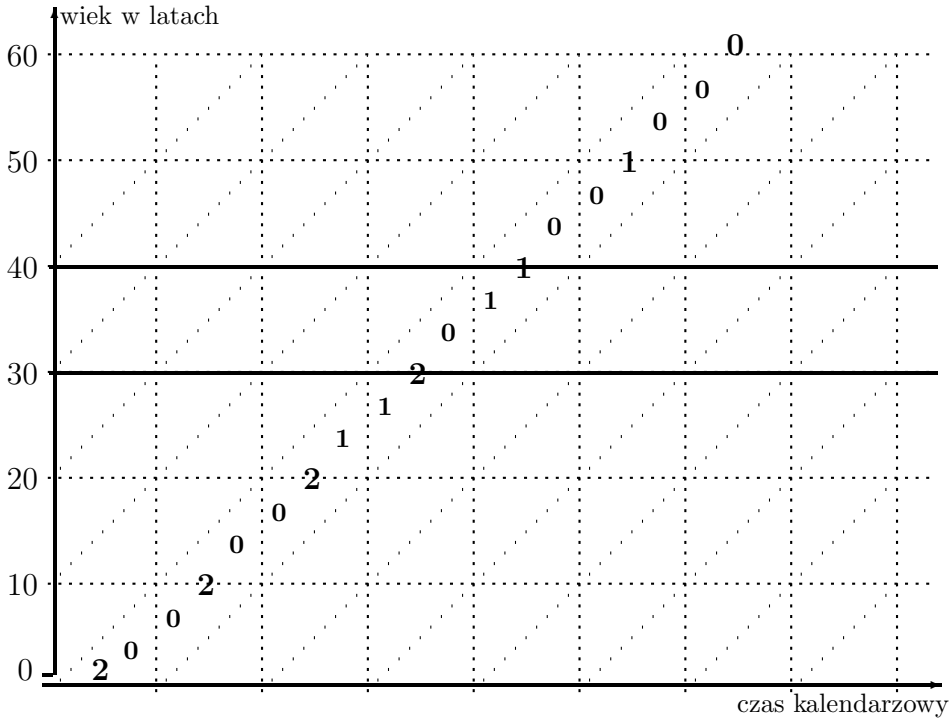
Z kolei cząstkowy współczynnik zgonów obliczymy na podstawie danych dotyczących wszystkich jednostek, tj. kobiet i mężczyzn łącznie. Oznaczmy przez ${}_{10}D_{30}$ liczbę zgonów, natomiast przez ${}_{10}K_{30}$ czas ekspozycji na ryzyko zgonu w danej grupie wieku. Mamy

$${}_{10}D_{30} = 1, \quad {}_{10}K_{30} = 10 + 10 + 10 + 8 = 38 \text{ osobołat.}$$

Otrzymujemy cząstkowy współczynnik zgonów równy

$${}_{10}m_{30} = \frac{1}{38}.$$

Następnie obliczymy wartości obu współczynników (płodności i zgonów) na podstawie danych zagregowanych. Rysunek 14 przedstawia fragment siatki demograficznej, a w niej liczby dożywających I rodzaju (kobiet) i liczby urodzonych dzieci (odpowiednio na odcinkach poziomych oraz wewnątrz trójkątnych pól).



Rysunek 14. Fragmenty siatek demograficznych, liczby dożywających I rodzaju, wybrany przedział wieku i liczby urodzeń (na podstawie rys. 13)

Na tej podstawie określimy liczbę urodzeń ${}_{10}B_{30}$ oraz przybliżymy czas ekspozycji ${}_{10}K_{30}^{(fem)}$ na szansę urodzenia dziecka dla kobiet w grupie wieku $[30, 40)$ lat. Mamy

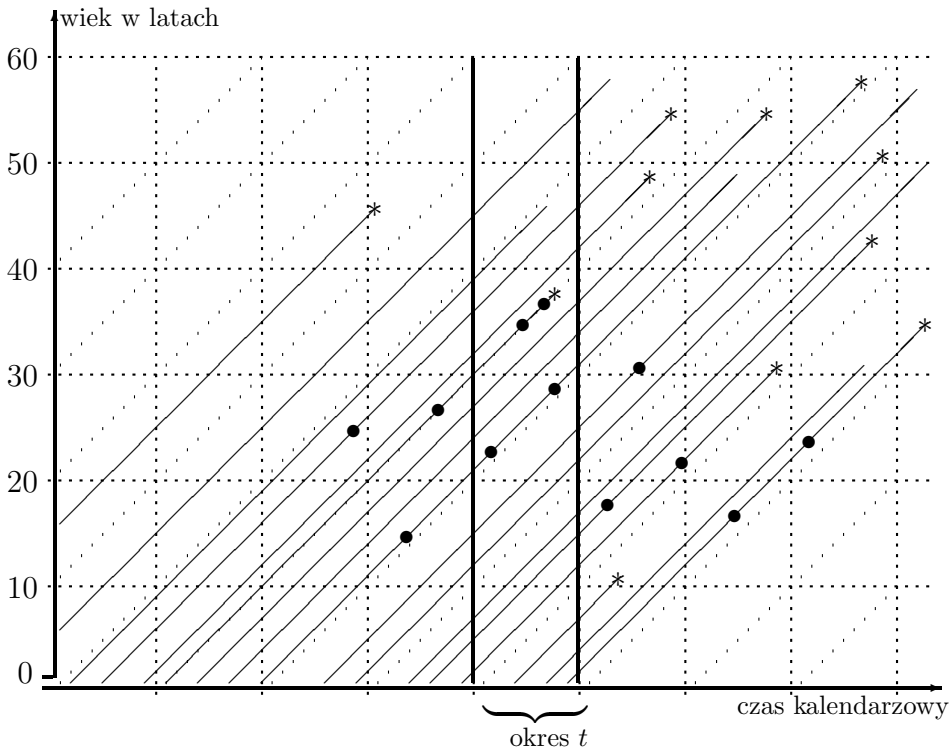
$${}_{10}B_{30} = 1, \quad {}_{10}K_{30}^{(fem)} \approx \frac{2+1}{2} \cdot 10 = 15 \text{ osobolat.}$$

Stąd cząstkowy współczynnik płodności jest równy w przybliżeniu

$${}_{10}f_{30} \approx \frac{1}{15}.$$

Następna siatka (rysunek 15) zawiera dane zagregowane dotyczące liczb dożywających I rodzaju (kobiet i mężczyzn) oraz liczb zgonów w analizowanej kohorcie. Na tej podstawie znajdziemy wartości ${}_{10}D_{30}$ oraz ${}_{10}K_{30}$ dla członków kohorty będących w grupie wieku $[30, 40)$ lat. Mamy w tym przypadku

$${}_{10}D_{30} = 1, \quad {}_{10}K_{30} \approx \frac{4+3}{2} \cdot 10 = 35 \text{ osobolat.}$$



Rysunek 16. Fragment siatki demograficznej, przykładowe linie życia kilku generacji oraz wybrany przedział czasu kalendarzowego t

Liczba zgonów D oraz liczba urodzeń B w okresie t równe są w tym przykładzie odpowiednio

$$D = 2, \quad B = 4.$$

Czas ekspozycji K na ryzyko zgonu w okresie t dla ogółu jednostek jest równy

$$K = 11 \cdot 10 + 6 + 8 + 6 + 2 + 1 = 133 \text{ osobołat},$$

natomiast czas ekspozycji kobiet w wieku rozrodczym na szansę urodzenia dziecka w tym okresie wynosi

$$K^{(fem)} = 4 \cdot 10 + 8 + 5 = 53 \text{ osobołat}.$$

Otrzymujemy ogólne współczynniki zgonów i płodności

$$CDR = \frac{2}{133}, \quad GFR = \frac{4}{53}.$$

Następnie wartości CDR i GFR przybliżymy, korzystając z liczb dożywających II rodzaju, które są określone przez liczby linii życia przecinających początek i koniec rozważanego okresu. Wynoszą one odpowiednio 13 oraz 14, a dla kobiet w wieku rozrodczym – są równe w obu przypadkach 5. Na tej podstawie szacujemy czas ekspozycji K na ryzyko zgonu w okresie t dla danej zbiorowości i czas ekspozycji $K^{(fem)}$ kobiet na szansę urodzenia dziecka w tym okresie, czyli

$$K \approx \frac{13 + 14}{2} \cdot 10 = 135 \text{ osobolat}, \quad K^{(fem)} \approx \frac{5 + 5}{2} \cdot 10 = 50 \text{ osobolat}.$$

Wobec tego współczynniki CDR i GFR są w przybliżeniu równe

$$CDR \approx \frac{2}{135}, \quad GFR \approx \frac{4}{50}.$$

W celu obliczenia cząstkowego współczynnika zgonów w przedziale wieku $[30, 40)$ lat oraz cząstkowego współczynnika płodności kobiet w tej samej grupie wieku, znajdujemy liczbę zgonów, liczbę urodzeń, czas ekspozycji osób będących w wieku $[30, 40)$ lat na ryzyko zgonu w okresie t , a także czas ekspozycji kobiet na szansę urodzenia dziecka w analogicznej grupie wieku i okresie. Oznaczmy te wielkości odpowiednio ${}_{10}D_{30}$, ${}_{10}B_{30}$, ${}_{10}K_{30}$, ${}_{10}K_{30}^{(fem)}$. Mamy na podstawie linii życia z rysunku 16

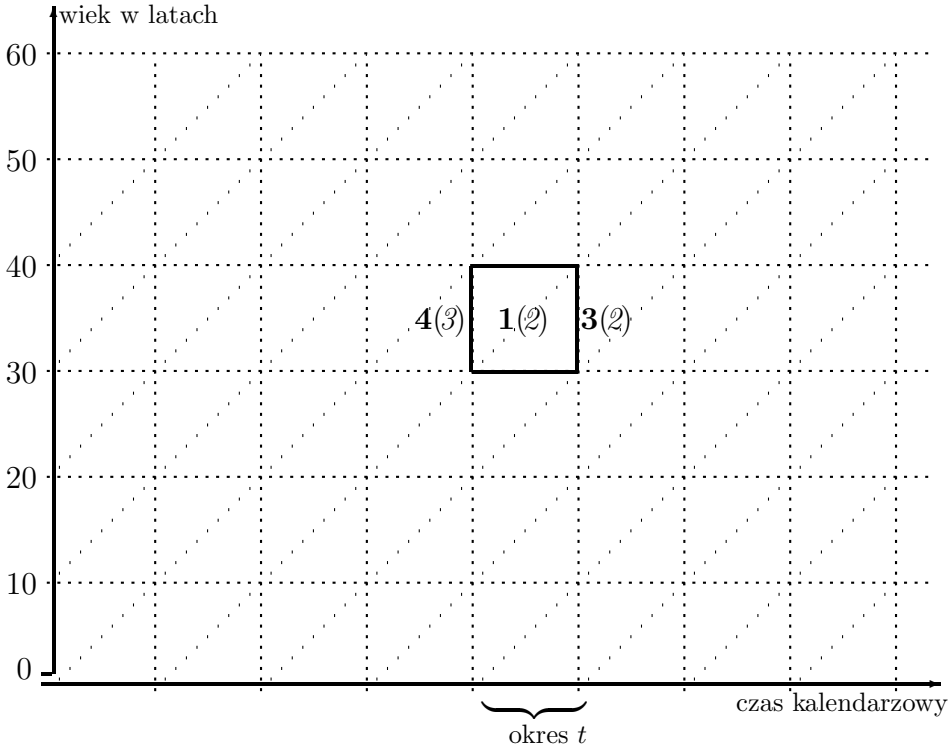
$${}_{10}D_{30} = 1, \quad {}_{10}K_{30} = 1 + 4 + 8 + 8 + 7 + 1 + 4 = 33 \text{ osobolata},$$

$${}_{10}B_{30} = 2, \quad {}_{10}K_{30}^{(fem)} = 4 + 8 + 8 + 7 + 1 = 28 \text{ osobolat}.$$

Na tej podstawie otrzymujemy

$${}_{10}m_{30} = \frac{1}{33}, \quad {}_{10}f_{30} = \frac{2}{28}.$$

Wartości przybliżone obu współczynników cząstkowych uzyskamy na podstawie liczb zdarzeń III rodzaju, tj. urodzeń lub zgonów, a także liczb dożywających II rodzaju. Na rysunku 17 liczby zgonów i urodzeń zamieszczono wewnątrz kwadratu korespondującego z wybraną grupą wieku $[30, 40)$ lat i okresem kalendarzowym t (liczba zgonów – czcionka pogrubiona, liczba urodzeń – czcionka pochylona). Z kolei liczby z lewej i prawej strony kwadratu dotyczą dożywających II rodzaju do początku i końca okresu t (w tym ogółu badanych – czcionka pogrubiona, kobiet – czcionka pochylona).



Rysunek 17. Fragment siatki demograficznej, liczby dożywających II rodzaju i zdarzeń III rodzaju: zgony/urodzenia (na podstawie rysunku 16)

Mamy zatem

$${}_{10}D_{30} = 1,$$

$${}_{10}K_{30} \approx \frac{4+3}{2} \cdot 10 = 35 \text{ osobo\l at}$$

oraz

$${}_{10}B_{30} = 2,$$

$${}_{10}K_{30}^{(fem)} \approx \frac{3+2}{2} \cdot 10 = 25 \text{ osobo\l at},$$

z czego dostajemy przybliżenia obu współczynników cząstkowych

$${}_{10}m_{30} \approx \frac{1}{35}, \quad {}_{10}f_{30} \approx \frac{2}{25}.$$

2.8. Prawdopodobieństwa zdarzeń demograficznych – ujęcie kohortowe i kohortowo-przekrojowe

W przypadku prawdopodobieństw zdarzeń demograficznych korzysta się na ogół z klasycznej definicji prawdopodobieństwa.

Załóżmy, że zdefiniowany jest pewien n -elementowy zbiór zdarzeń elementarnych Ω , jednakowo prawdopodobnych. W demografii Ω reprezentuje zbiorowość osób lub grup osób (np. małżeństw, rodzin, gospodarstw domowych).

Zgodnie z definicją klasyczną, prawdopodobieństwo p wystąpienia określonego zdarzenia obliczamy jako

$$p = \frac{k}{n}, \quad (2.38)$$

gdzie k – liczba zdarzeń określonego rodzaju, które wystąpiły w danej zbiorowości jednostek, n – łączna liczebność zbiorowości.

W analizie demograficznej prawdopodobieństwa zdarzeń obliczamy w ujęciu kohortowym lub kohortowo-przekrojowym.

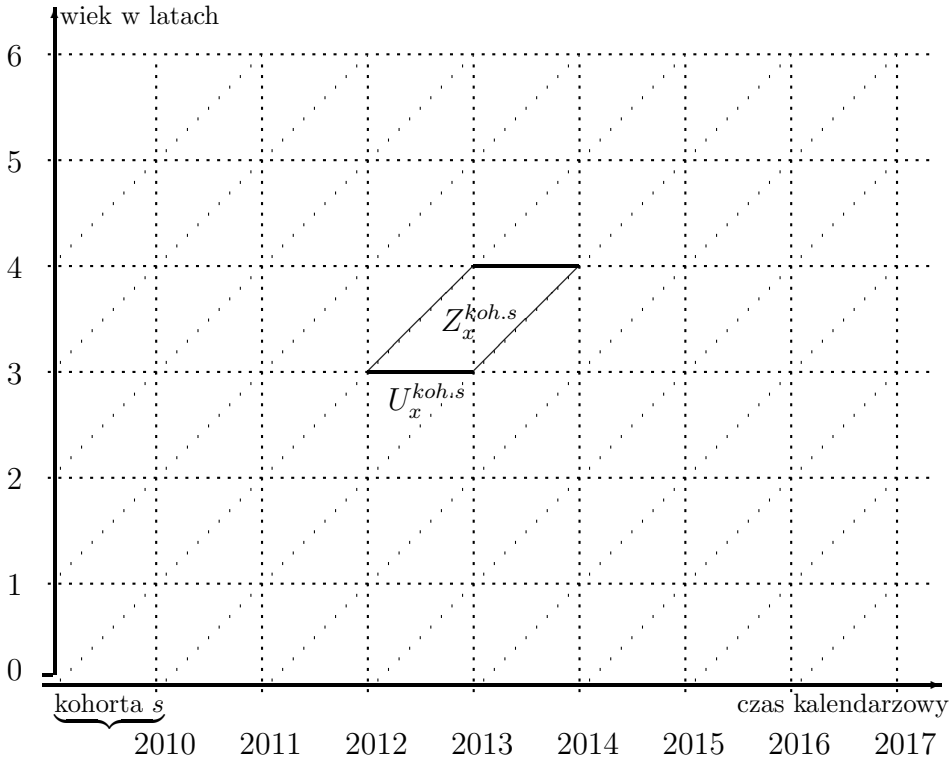
W **analizie kohortowej**, jeśli dysponujemy danymi o liczbie zdarzeń I rodzaju $Z_x^{koh.s}$ w rocznych grupach wieku oraz danymi o liczbie dożywających I rodzaju $U_x^{koh.s}$ w ustalonej kohorcie s (rysunek 18), prawdopodobieństwo wystąpienia interesującego zdarzenia w tej grupie wieku definiujemy jako iloraz

$$q_x^{koh.s} = \frac{Z_x^{koh.s}}{U_x^{koh.s}}, \quad (2.39)$$

a pomijając oznaczenie numeru kohorty, w postaci

$$q_x = \frac{Z_x}{U_x}. \quad (2.40)$$

Dla porównania, kohortowy cząstkowy współczynnik demograficzny ma postać (2.2), tj. $W_x^{koh.s} = Z_x^{koh.s} / K_x^{koh.s}$, gdzie $K_x^{koh.s}$ jest łącznym czasem ekspozycji osób w wieku x ukończonych lat na ryzyko wystąpienia danego zdarzenia.

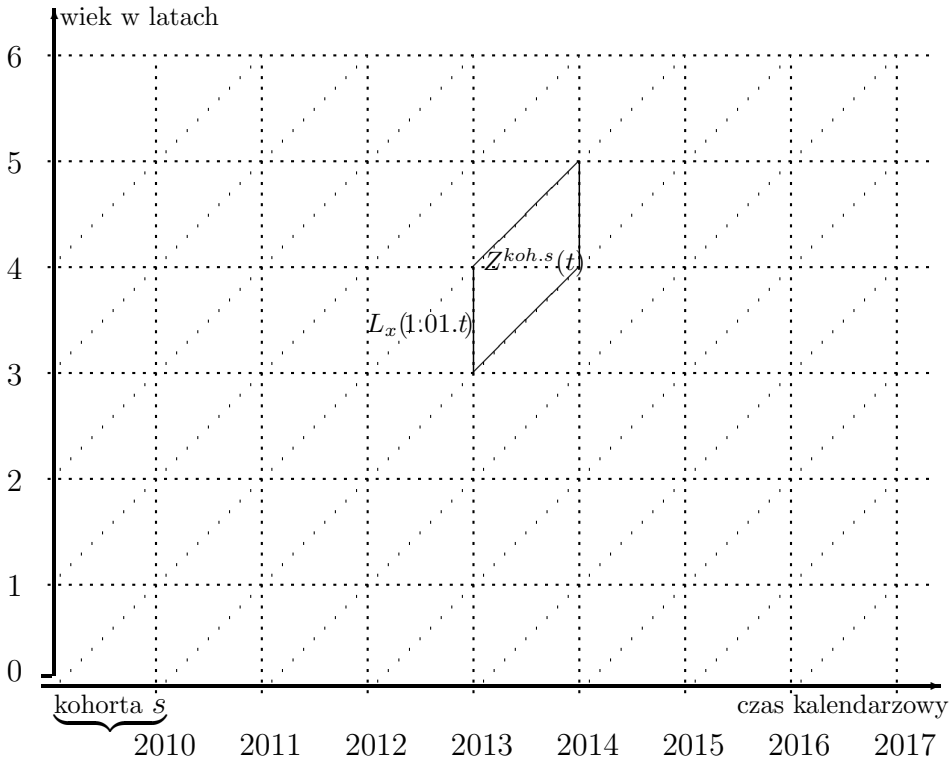


Rysunek 18. Fragment siatki demograficznej, liczby dożywających I rodzaju oraz liczby zdarzeń I rodzaju

Podobnie w **analizie kohortowo-przekrojowej**, jeśli dysponujemy danymi o liczbie zdarzeń II rodzaju oraz o liczbie dożywających II rodzaju (rysunek 19), wówczas prawdopodobieństwo wystąpienia interesującego nas zdarzenia w roku t i w badanej kohorcie obliczymy ze wzoru

$$q^{persp.}(t) = \frac{Z_x^{koh.s}(t)}{L_x(1.01.t)}. \quad (2.41)$$

Dla przypomnienia, kohortowo-przekrojowy współczynnik demograficzny dla roku t i zadanej kohorty s wyraża się wzorem (2.12), tj. $W_x^{koh.s}(t) = Z_x^{koh.s}(t)/K_x^{koh.s}(t)$, gdzie $K_x^{koh.s}(t)$ jest czasem ekspozycji w roku t dla jednostek kohorty będących w wieku x ukończonych lat. Jak łatwo zauważyć, różnice występują w mianownikach obu formuł.



Rysunek 19. Fragment siatki demograficznej, liczby dożywających II rodzaju oraz liczby zdarzeń II rodzaju

Prawdopodobieństwa q_x oraz $q^{persp}(t)$ różnią się z punktu widzenia zarówno konstrukcji, jak i interpretacji:

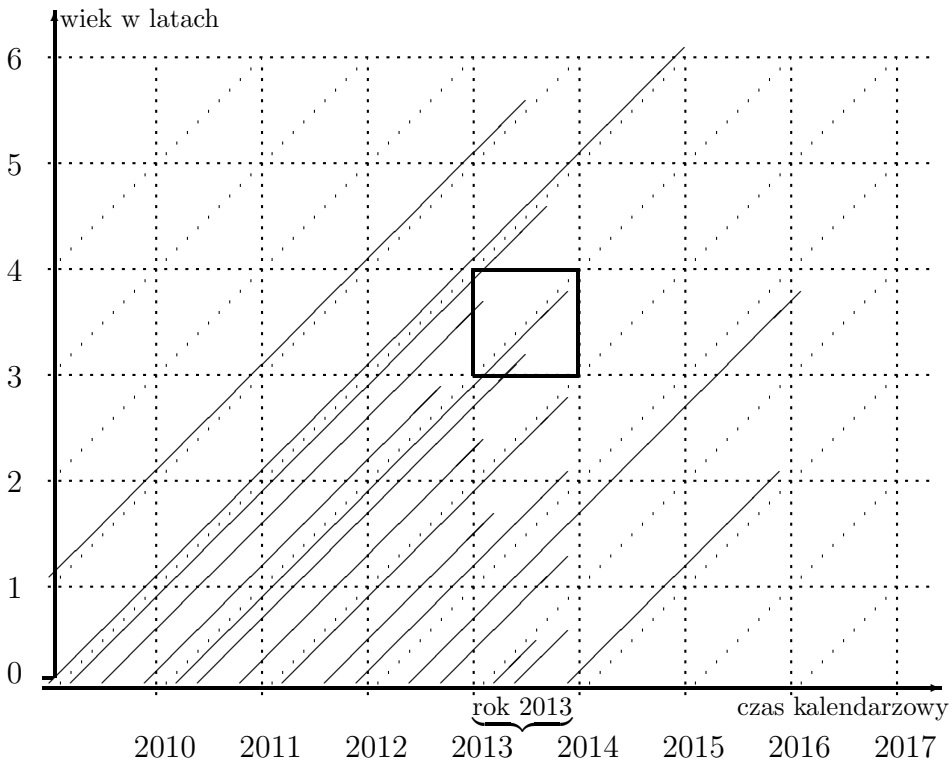
q_x – oznacza prawdopodobieństwo wystąpienia określonego zdarzenia w ciągu roku dla tych jednostek kohorty, które dożyły określonego momentu czasu własnego (np. do x -tych urodzin),

$q^{persp}(t)$ – oznacza prawdopodobieństwo wystąpienia zdarzenia w ciągu roku dla tych jednostek kohorty, które dożyły określonego momentu czasu kalendarzowego (np. do 1 stycznia roku t). Prawdopodobieństwo to nazywamy perspektywicznym.

Należy zauważyć, że w przypadku **danych przekrojowych** nie wyznaczamy prawdopodobieństw zdarzeń demograficznych jako ilorazów liczb zdarzeń III rodzaju oraz liczb dożywających II rodzaju. Na przykład, nie obliczamy prawdopodobieństw zgonu, dzieląc łączną liczbę zgonów w roku t w grupie osób w wieku x ukończonych lat przez ogólną liczbę osób w tym wieku dożywających do początku roku kalendarzowego t . Problem ten jest wyjaśniony bliżej w przykładzie 7.

Przykład 7. Rozważmy umowną zbiorowość dzieci, których linie życia ilustruje rysunek 20.

Liczba zgonów trzylatków w roku 2013 wynosi $D_3 = 3$, ale liczba trzylatków, która dożyła do dnia 1 stycznia 2013 roku, wynosi tylko $L_3(1.01.13) = 2$, bowiem tylko dwie linie życia przebiegają przez przedział wieku $[3, 4)$ lat i przecinają jednocześnie pionową linię siatki wyznaczającą początek roku 2013.



Rysunek 20. Fragment siatki demograficznej, przykładowe linie życia oraz wybrane przedziały czasu kalendarzowego (rok 2013) i wieku $[3, 4)$ lata

Iloraz

$$\frac{D_3}{L_3(1.01.t)} = \frac{3}{2}$$

nie może jednak reprezentować prawdopodobieństwa zgonu w grupie trzylatków, ponieważ otrzymany wynik jest większy od 1.

Przykład ten pokazuje, że dane przekrojowe w ogólnym przypadku nie nadają się do obliczania prawdopodobieństw zdarzeń, ponieważ liczby zdarzeń III rodzaju i liczby dożywających II rodzaju nie odnoszą się do tych samych zbiorowości.

2.9. Związek współczynników demograficznych i prawdopodobieństw zdarzeń

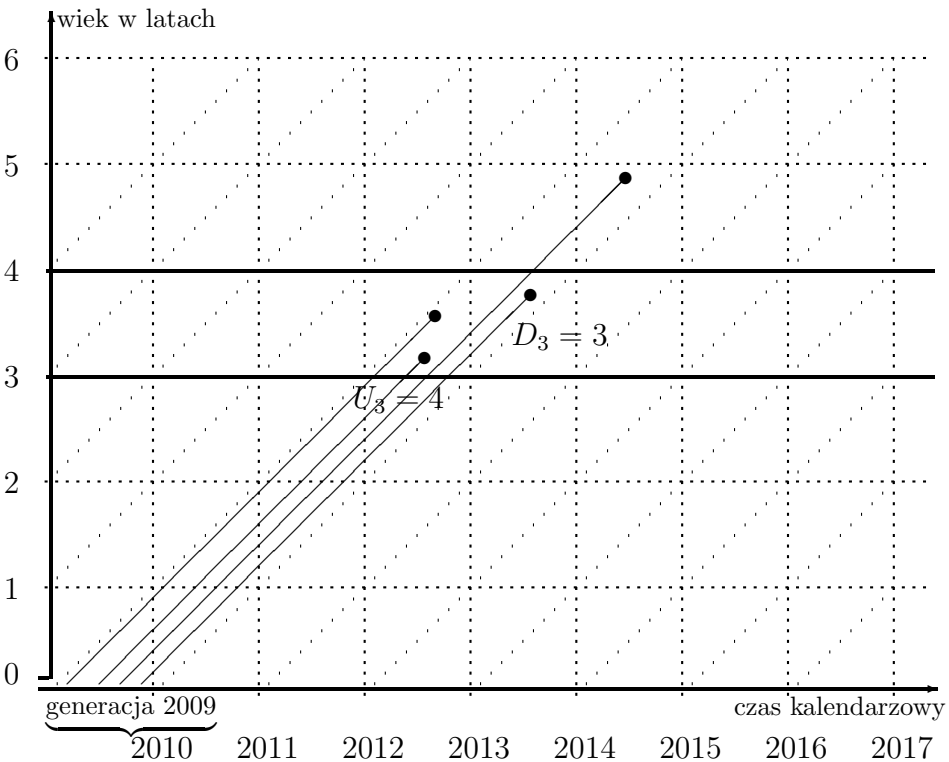
Relację wiążącą współczynnik demograficzny i prawdopodobieństwo zdarzeń przedstawimy na przykładzie kohortowego, cząstkowego współczynnika zgonów $m_x^{koh.s}$ i prawdopodobieństwa zgonu $q_x^{koh.s}$.

Na wstępie rozważmy przykład wyjaśniający związek

$$K_x = U_x - D_x + D_x \cdot a_x, \quad (2.42)$$

gdzie K_x oznacza czas ekspozycji osób w wieku x ukończonych lat, a_x oznacza średni czas życia w ciągu roku tych x -latków, którzy nie dożyli $x + 1$ urodzin, U_x jest liczbą dożywających wieku x , a D_x jest liczbą zgonów w grupie wieku $[x, x + 1)$ lat.

Przykład 8. Niech dana będzie kohorta urodzeniowa, w której linie życia ilustruje rysunek 21. Załóżmy, że badanymi zdarzeniami w tej kohorcie są zgony w grupie wieku $x = 3$ ukończonych lat.



Rysunek 21. Fragment siatki demograficznej, przykładowe linie życia generacji urodzonych w 2009 roku oraz wybrany przedział wieku $[3, 4)$ lata

Liczbę zgonów w grupie wieku $[3, 4)$ lat oznaczmy symbolem D_3 . Wówczas czas ekspozycji K_3 na ryzyko zgonu w analizowanej grupie wieku można zapisać następująco

$$\begin{aligned} K_3 &= \{\text{długość przedz. wieku } [3,4)\} \cdot \{\text{liczba dożywających 4 lat}\} \\ &+ \{\text{liczba zgonów wśród trzylatków}\} \cdot \{\text{średni czas życia trzylatków,} \\ &\quad \text{którzy zmarli, nie dożywając czwartych urodzin}\} = \\ &= 1 \cdot (4 - 3) + 3 \cdot \frac{0,6 + 0,2 + 0,8}{3} = 1 + 3 \cdot 0,53 = 2,59, \end{aligned}$$

Powyższe równanie można zapisać symbolami jako

$$K_3 = n \cdot (U_3 - D_3) + D_3 \cdot a_3,$$

gdzie $U_3 = 4$ to liczba dożywających I rodzaju do wieku 3 lat, $D_3 = 3$ jest liczbą zdarzeń demograficznych (zgonów) w generacji trzylatków, $a_3 = 0,53$ jest średnim czasem życia trzylatków zmarłych w ciągu roku, natomiast $n = 1$ to długość przedziału wieku $[3, 4)$ lat.

Uogólniając rozważania z przykładu 8 na przypadek przedziałów $[x, x + n)$ dla $n > 1$ ($n \in \mathbf{N}$), możemy zapisać

$${}_nK_x = n \cdot (U_x - {}_nD_x) + {}_nD_x \cdot {}_na_x, \quad (2.43)$$

co daje

$$n \cdot U_x = {}_nK_x + n \cdot {}_nD_x - {}_nD_x \cdot {}_na_x = {}_nK_x + (n - {}_na_x) \cdot {}_nD_x, \quad (2.44)$$

stąd

$$\begin{aligned} {}_nq_x &= \frac{{}_nD_x}{U_x} = \frac{n \cdot {}_nD_x}{{}_nK_x + (n - {}_na_x) \cdot {}_nD_x} = \\ &= \frac{n \cdot \frac{{}_nD_x}{{}_nK_x}}{\frac{{}_nK_x}{{}_nK_x} + (n - {}_na_x) \cdot \frac{{}_nD_x}{{}_nK_x}} = \frac{n \cdot {}_nm_x}{1 + (n - {}_na_x) \cdot {}_nm_x}. \end{aligned} \quad (2.45)$$

Ostatecznie otrzymujemy

$${}_nq_x = \frac{n \cdot {}_nm_x}{1 + (n - {}_na_x) \cdot {}_nm_x}. \quad (2.46)$$

Wzór (2.46) przedstawia relację wiążącą kohortowy współczynnik zgonów ${}_nm_x$ z prawdopodobieństwem zgonu ${}_nq_x$.

Warto wspomnieć, że choć prawdopodobieństw zgonu nie wyznaczamy z danych przekrojowych, to jednak możemy je oszacować na podstawie relacji (2.46), wykorzystując przekrojowe, cząstkowe współczynniki zgonów.

Przypadek szczególny: Rozważmy roczne grupy wieku $[x, x + 1)$. Wówczas (2.43) upraszcza się do zapisu

$$K_x = U_x - D_x + D_x \cdot a_x, \quad (2.47)$$

natomiast relacja (2.46) redukuje się do postaci

$$q_x = \frac{m_x}{1 + (1 - a_x) \cdot m_x}. \quad (2.48)$$

Przyjmijmy często stosowane przybliżenie $a_x \approx \frac{1}{2}$. Należy zaznaczyć, że przybliżenie to nie jest adekwatne dla grupy wieku $x = 0$ lat. Wśród niemowląt umieralność jest wysoka w pierwszych tygodniach po urodzeniu, a więc współczynnik a_0 kształtuje się w tej grupie na poziomie znacznie poniżej 0,5.

Mamy dla $x = 1, \dots, X$

$$K_x \approx U_x - \frac{1}{2}D_x \quad (2.49)$$

oraz

$$q_x = \frac{m_x}{1 + \frac{1}{2} \cdot m_x} = \frac{2m_x}{2 + m_x}. \quad (2.50)$$

Ponieważ

$$D_x = U_x - U_{x+1}, \quad (2.51)$$

więc otrzymujemy również

$$K_x \approx U_x - \frac{1}{2}(U_x - U_{x+1}) = \frac{1}{2}(U_x + U_{x+1}). \quad (2.52)$$

Stąd wynika często stosowana metoda przybliżania współczynnika zgonów za pomocą wyrażenia

$$m_x = \frac{D_x}{K_x} \approx \frac{D_x}{\frac{1}{2}(U_x + U_{x+1})}. \quad (2.53)$$

Zauważymy także, że dla współczynnika (2.53), spełniona jest nadal relacja (2.50), czyli

$$\begin{aligned} \frac{2m_x}{2 + m_x} &= \frac{2D_x}{\frac{1}{2}(U_x + U_{x+1}) \left(2 + \frac{D_x}{\frac{1}{2}(U_x + U_{x+1})}\right)} = \\ &= \frac{2D_x}{(U_x + U_{x+1}) \frac{U_x + U_{x+1} + D_x}{U_x + U_{x+1}}} = \frac{2D_x}{2U_x} = q_x. \end{aligned} \quad (2.54)$$

2.10. Standaryzacja i dekompozycja współczynników demograficznych

Współczynniki demograficzne służą do oceny natężenia i poziomu zjawisk demograficznych. W wielu przypadkach zachodzi również potrzeba ich porównania w czasie, w przestrzeni lub pomiędzy subpopulacjami.

Porównanie w czasie pozwala określać kierunek zmian, porównanie w przestrzeni – ustalić miejsce badanego regionu w relacji do innych regionów, porównanie pomiędzy subpopulacjami – ocenić różnice w poziomie zjawiska w poszczególnych subpopulacjach.

W analizach porównawczych dotyczących zjawisk demograficznych użyteczne jest korzystanie z metod **standaryzacji i dekompozycji** współczynników.

Przybliżymy to zagadnienie na przykładzie ogólnego współczynnika, oznaczonego dalej symbolem $W(t)$, gdzie t oznaczać będzie okres, region lub subpopulację. Mamy z definicji

$$W(t) = \frac{Z(t)}{K(t)} \cdot C, \quad (2.55)$$

gdzie:

$Z(t)$ – liczba zdarzeń w okresie, regionie lub subpopulacji t ,

$K(t)$ – liczba osobolat życia (czasu ekspozycji) jednostek w okresie, regionie lub subpopulacji t ,

C – ustalona stała.

Niech $x = 0, 1, \dots, X$ będą indeksami identyfikującymi kategorie pewnej cechy społeczno-ekonomicznej (np. wieku, płci, wykształcenia

itp.), natomiast $W_x(t)$ niech oznacza dowolny cząstkowy współczynnik demograficzny w x -tej podzbiorowości w okresie, regionie lub subpopulacji t , wyrażony przez analogię do (2.55) jako iloraz

$$W_x(t) = \frac{Z_x(t)}{K_x(t)} \cdot C. \quad (2.56)$$

Przy tych oznaczeniach (2.55) można przedstawić w postaci

$$W(t) = \frac{\sum_{x=0}^X Z_x(t)}{\sum_{x=0}^X K_x(t)} \cdot C. \quad (2.57)$$

Po kilku przekształceniach otrzymujemy

$$W(t) = \frac{\sum_{x=0}^X \frac{Z_x(t)}{K_x(t)} K_x(t)}{\sum_{x=0}^X K_x(t)} \cdot C = \frac{\sum_{x=0}^X W_x(t) K_x(t)}{\sum_{x=0}^X K_x(t)} = \quad (2.58)$$

$$= \sum_{x=0}^X W_x(t) \frac{K_x(t)}{\sum_{x=0}^X K_x(t)}.$$

Oznaczając

$$k_x(t) = \frac{K_x(t)}{\sum_{x=0}^X K_x(t)}, \quad (2.59)$$

mamy także

$$W(t) = \sum_{x=0}^X W_x(t) k_x(t). \quad (2.60)$$

Z zapisu (2.60) wynika, że ogólny współczynnik demograficzny $W(t)$ jest ważoną sumą współczynników cząstkowych $W_x(t)$ z wagami $k_x(t)$ reprezentującymi udziały osobolat życia (czasów ekspozycji) podzbiorowości x w łącznej liczbie osobolat życia jednostek populacji w okresie, regionie lub subpopulacji t .

W praktyce wagi $k_x(t)$ przybliżamy za pomocą wzoru

$$k_x(t) \approx \frac{\bar{L}_x(t)}{\sum_{x=0}^X \bar{L}_x(t)},$$

gdzie $\bar{L}_x(t)$ jest średnią liczebnością x -tej podzbiorowości w okresie, regionie lub subpopulacji t .

Uwaga 11: Jeśli t oznacza okres kalendarzowy (np. ustalony rok), to w miejsce $\bar{L}_x(t)$ przyjmujemy na ogół średnią ze stanu liczebnego podziorowości x na początku i końcu danego okresu lub liczebność w połowie roku, tj. w dniu 30 czerwca.

Przykład 9. Rozważmy liczbę zgonów w roku 2017 oraz liczbę ludności w Polsce w połowie roku 2017, w podziale wg 10-letnich grup wieku (tablica 1).

Obliczymy cząstkowe współczynniki zgonów, wskaźniki struktury ludności wg wieku, a następnie ogólny współczynnik zgonów dla 2017 roku.

Tablica 1. Liczba zgonów oraz stan ludności
na dzień 30 czerwca 2017 roku

Grupy wieku	Liczba zgonów	Ludność (30.06.17)
(1)	(2)	(3)
0 – 10	2023	3 961 035
10 – 20	870	3 730 509
20 – 30	3087	4 976 155
30 – 40	6866	6 310 291
40 – 50	13 968	5 272 064
50 – 60	36 726	5 007 944
60 – 70	81 625	5 074 712
70 – 80	82 816	2 469 250
80+	174 871	1 620 386
Razem	402 852	38 422 346

Źródło: dane GUS, stat.gov.pl

Rozwiązanie

Cząstkowe współczynniki zgonów w poszczególnych grupach wieku znajdziemy, dzieląc przez siebie wyrazy z kolumn 2 i 3 tablicy 1. Rezultaty dzielenia pomnożymy dodatkowo przez stałą $C = 1000$. Wyniki obliczeń zawiera kolumna 5 tablicy 2.

Z kolei wskaźniki struktury wyznaczymy na podstawie danych z kolumny 3 tablicy 1, jako ilorazy liczby ludności w poszczególnych grupach wieku do ogólnej liczebności. Wyniki przedstawia kolumna 4 tablicy 2.

W celu wyznaczenia **ogólnego współczynnika zgonów** zastosujemy dwie metody. W pierwszej wykorzystamy definicję (2.29) współczynnika *CDR*, jako ilorazu ogólnej liczby zgonów i średniej liczby ludności w danym roku. W drugiej metodzie wykorzystamy sumę ważoną współczynników cząstkowych (2.60).

Metoda 1 – z definicji współczynnika *CDR*:

$$\begin{aligned} CDR &= \frac{\text{liczba zgonów w roku 2017}}{\text{liczba ludności w połowie roku 2017}} \cdot \text{stała} = \\ &= \frac{402\,852}{38\,422\,346} \cdot 1000 \approx 10,485. \end{aligned}$$

Metoda 2 – suma ważona współczynników cząstkowych:

$$\begin{aligned} CDR &= 0,511 \cdot 0,103 + 0,233 \cdot 0,097 + 0,620 \cdot 0,130 + \\ &+ 1,088 \cdot 0,164 + 2,649 \cdot 0,137 + 7,334 \cdot 0,130 + \\ &+ 16,085 \cdot 0,132 + 33,539 \cdot 0,065 + 107,919 \cdot 0,042 \approx 10,485. \end{aligned}$$

Tablica 2. Liczba zgonów, stan ludności na 30 czerwca 2017 roku, struktura wg wieku i współczynniki zgonów wg grup wieku

Grupy wieku	Liczba zgonów	Ludność (30.06.17)	Wskaźniki struktury	Wsp. zgonów na 1000 osób
(1)	(2)	(3)	(4)	(5)
0 – 10	2023	3 961 035	0,103	0,511
10 – 20	870	3 730 509	0,097	0,233
20 – 30	3087	4 976 155	0,130	0,620
30 – 40	6866	6 310 291	0,164	1,088
40 – 50	13 968	5 272 064	0,137	2,649
50 – 60	36 726	5 007 944	0,130	7,334
60 – 70	81 625	5 074 712	0,132	16,085
70 – 80	82 816	2 469 250	0,065	33,539
80+	174 871	1 620 386	0,042	107,919
Razem	402 852	38 422 346	1,000	×

Źródło: dane GUS, obliczenia własne

Jak łatwo zauważyć, obie metody prowadzą do takiej samej wartości współczynnika CDR , tj. do liczby 10, 485. Wielkość ta określa średnią liczbę zgonów w Polsce w roku 2017 przypadającą na 1000 ludności.

Druga metoda użyta w tym przykładzie, oparta na iloczynie współczynników cząstkowych i wskaźników struktury, pokazuje dodatkowo, że na poziom ogólnego współczynnika CDR ma wpływ zarówno natężenie umieralności w poszczególnych grupach wieku, jak również struktura ludności wg wieku.

W podsumowaniu możemy stwierdzić, że o wartości ogólnego współczynnika demograficznego, tj. sumy ważonej współczynników cząstkowych, decydują dwa główne czynniki:

- 1) natężenie zjawiska w podzbiorowościach x , reprezentowane przez współczynniki cząstkowe $W_x(t)$ w badanym okresie, regionie lub subpopulacji t ,
- 2) struktura zbiorowości, reprezentowana przez udziały $k_x(t)$ poszczególnych podzbiorowości w łącznej liczbie osobolat życia (czasie ekspozycji) w okresie, regionie lub subpopulacji t .

W ogólnym przypadku może się zdarzyć, że w porównywanych okresach, regionach lub subpopulacjach współczynniki cząstkowe są podobne, a mimo to wartości współczynnika ogólnego wykazują znaczne różnice lub na odwrót. Wskazane jest wówczas zbadanie, jaki wpływ na wartość tego współczynnika ma jeden z wyżej wymienionych czynników, przy niezmiennym poziomie drugiego z nich. W tym celu pomocne są tzw. **współczynniki standaryzowane**.

Jeśli w formule (2.60) wartości wag $k_x(t)$ zastąpimy pewnymi hipotetycznymi wagami $k_x^{(hip)}$ lub wartości współczynników cząstkowych $W_x(t)$ zastąpimy hipotetycznymi współczynnikami $W_x^{(hip)}$ wziętymi na przykład z ustalonego okresu referencyjnego (lub ustalonego regionu, subpopulacji), to w efekcie otrzymamy dwa standaryzowane współczynniki demograficzne postaci

$$W^{(b)}(t) = \sum_{x=0}^X W_x(t)k_x^{(hip)}, \quad (2.61)$$

$$W^{(p)}(t) = \sum_{x=0}^X W_x^{(hip)}k_x(t). \quad (2.62)$$

Formuły (2.61), (2.62) definiują współczynniki standaryzowane metodą **bezpośrednią** i **pośrednią**. Służą one za podstawę porównań różnych okresów, regionów lub subpopulacji.

Co więcej, jeśli w miejsce hipotetycznych współczynników $W_x^{(hip)}$ lub wskaźników struktury $k_x^{(hip)}$ przyjmujemy wartości średnie z dwóch porównywanych okresów, regionów lub subpopulacji t_1 i t_2 , to w kolejnym kroku możemy dokonać prostej **dekompozycji** różnicy współczynników $W(t_1) - W(t_2)$ na sumę dwóch różnic postaci

$$W(t_1) - W(t_2) = [W^{(b)}(t_1) - W^{(b)}(t_2)] + [W^{(p)}(t_1) - W^{(p)}(t_2)]. \quad (2.63)$$

Pierwszy składnik $W^{(b)}(t_1) - W^{(b)}(t_2)$ po prawej stronie (2.63) interpretujemy jako efekt różnic w natężeniu zjawiska, a drugi składnik $W^{(p)}(t_1) - W^{(p)}(t_2)$ jako efekt różnic w strukturze obu porównywanych zbiorowości.

Przykład 10. Na podstawie danych dotyczących cząstkowych współczynników zgonów w subpopulacji mężczyzn i kobiet w pewnej zbiorowości (tablice 3 i 4), oszacujemy ogólne współczynniki zgonów dla obu subpopulacji, a następnie współczynniki standaryzowane metodą bezpośrednią i pośrednią, przyjmując w pierwszym przypadku uśrednioną strukturę wg wieku, a w drugim – uśrednione cząstkowe współczynniki zgonów. Na tej podstawie dokonamy dekompozycji różnicy pomiędzy współczynnikami ogólnymi, korzystając z (2.63).

Rozwiązanie

Ogólne współczynniki zgonów dla obu subpopulacji znajdziemy na podstawie wzoru (2.60), tj. jako sumy ważone współczynników cząstkowych. Mamy na podstawie danych z tablic 3 i 4

$$CDR_M \approx 9,483,$$

$$CDR_K \approx 9,800,$$

co oznacza, że na 1000 mężczyzn przypada ok. 9,5 zgonów, natomiast na 1000 kobiet – ok. 9,8 zgonów.

Zauważymy, że współczynnik CDR dla mężczyzn jest nieco niższy niż dla kobiet, choć wszystkie współczynniki cząstkowe w subpopulacji mężczyzn są znacząco większe od analogicznych współczynników w subpopulacji kobiet (por. dane w kolumnie 2 w tablicach 3 i 4). Różnica obu współczynników ogólnych wynosi

$$CDR_M - CDR_K = -0,317.$$

Paradoks ten wynika z faktu, że struktura mężczyzn wg wieku różni się od analogicznej struktury kobiet. Większy udział wśród kobiet mają starsze grupy wieku. Oczywiście jest, że natężenie zgonów rośnie wraz z wiekiem, co wspólnie z wyższym udziałem starszych grup daje w efekcie większą wartość ogólnego współczynnika zgonów w subpopulacji kobiet.

Tablica 3. Częstkowe współczynniki zgonów oraz wskaźniki struktury subpopulacji mężczyzn wg 10-letnich grup wieku

Przedziały wieku	Współczynniki zgonów (na 1000 mężczyzn)	Wskaźniki struktury
(1)	(2)	(3)
0 – 10	0,563	0,109
10 – 20	0,302	0,103
20 – 30	0,974	0,136
30 – 40	1,657	0,172
40 – 50	3,872	0,143
50 – 60	10,641	0,132
60 – 70	23,142	0,125
70 – 80	46,686	0,075
80+	124,083	0,005

Źródło: dane GUS (częściowo zmienione), 2017

Tablica 4. Częstkowe współczynniki zgonów oraz wskaźniki struktury subpopulacji kobiet wg 10-letnich grup wieku

Przedziały wieku	Współczynniki zgonów (na 1000 kobiet)	Wskaźniki struktury
(1)	(2)	(3)
0 – 10	0,456	0,097
10 – 20	0,161	0,092
20 – 30	0,253	0,123
30 – 40	0,505	0,157
40 – 50	1,411	0,132
50 – 60	4,167	0,129
60 – 70	10,129	0,139
70 – 80	24,796	0,075
80+	100,703	0,056

Źródło: dane GUS, 2017

Aby wyeliminować wpływ struktury na wartości porównywanych współczynników, przyjmijmy dalej strukturę uśrednioną, czyli założymy hipotetyczne wskaźniki struktury w grupach wieku równe średnim wskaźnikom dla obu subpopulacji. W ten sposób otrzymamy standaryzowane współczynniki zgonów (przy stałej strukturze), oznaczone dalej jako $CDR_M^{(b)}$, $CDR_K^{(b)}$. Obliczając zgodnie ze wzorem (2.61), mamy

$$CDR_M^{(b)} \approx 12,748, \quad CDR_K^{(b)} \approx 7,184,$$

co daje różnicę równą

$$CDR_M^{(b)} - CDR_K^{(b)} = 5,564.$$

Rozważmy teraz współczynniki standaryzowane metodą pośrednią. Jeśli przyjmijmy, że natężenie zgonów, reprezentowane przez współczynniki cząstkowe, jest na poziomie średnim w obu porównywanych subpopulacjach, wówczas otrzymamy drugi zestaw współczynników standaryzowanych (przy stałym natężeniu), które oznaczymy symbolami $CDR_M^{(p)}$, $CDR_K^{(p)}$. Zgodnie z formułą (2.62), obydwie współczynniki są równe

$$CDR_M^{(p)} \approx 7,026, \quad CDR_K^{(p)} \approx 12,907,$$

a ich różnica wynosi

$$CDR_M^{(p)} - CDR_K^{(p)} = -5,881.$$

Zauważmy, że spełniona jest tu równość (2.63). Mamy bowiem

$$-0,317 = 5,564 - 5,881.$$

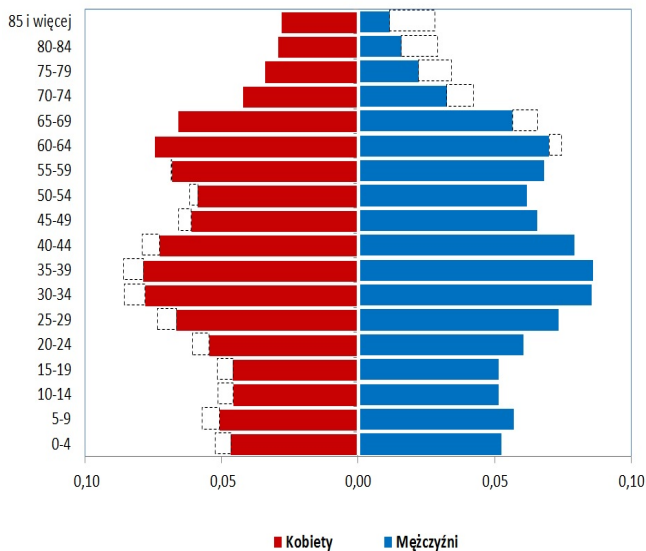
Z tego wynika, że na ujemną różnicę pomiędzy ogólnymi współczynnikami umieralności mężczyzn i kobiet, tj. $CDR_M - CDR_K = -0,317$, składa się efekt różnic w natężeniu umieralności cząstkowej (efekt rzędu 5,564), a także efekt odmiennych struktur obu porównywanych subpopulacji (na poziomie -5,881). Obydwie efekty co do wartości bezwzględnej są do siebie zbliżone, mają jednak odmiennie znaki, co wskazuje na przeciwny kierunek wpływu na różnicę pomiędzy umieralnością mężczyzn i kobiet.

Podsumowując, na ogólny poziom umieralności mężczyzn negatywnie wpływa przede wszystkim niekorzystny w porównaniu do kobiet

poziom natężenia umieralności w poszczególnych grupach wieku, czyli cząstkowych współczynników zgonów, co określane jest często mianem nadumieralności mężczyzn (ponad poziom umieralności kobiet).

Z drugiej strony, obserwujemy dodatni efekt struktury wg wieku subpopulacji mężczyzn w relacji do kobiet, będący w pewnym stopniu konsekwencją wspomnianej nadumieralności. W przypadku mężczyzn występują bowiem relatywnie większe niż u kobiet udziały młodszych grup wieku, a tym samym niższe udziały grup starszych.

Porównania struktury wg wieku subpopulacji mężczyzn w relacji do kobiet można dokonać stosunkowo łatwo na podstawie tzw. **piramidy wieku**, powstałej przez połączenie dwóch wykresów kolumnowych obrazujących obie subpopulacje w liczbach stosunkowych. Rysunek 22 ilustruje przykładową piramidę wieku dla populacji Polski z roku 2017.



Rysunek 22. Piramida wieku ludności w Polsce w 2017 roku

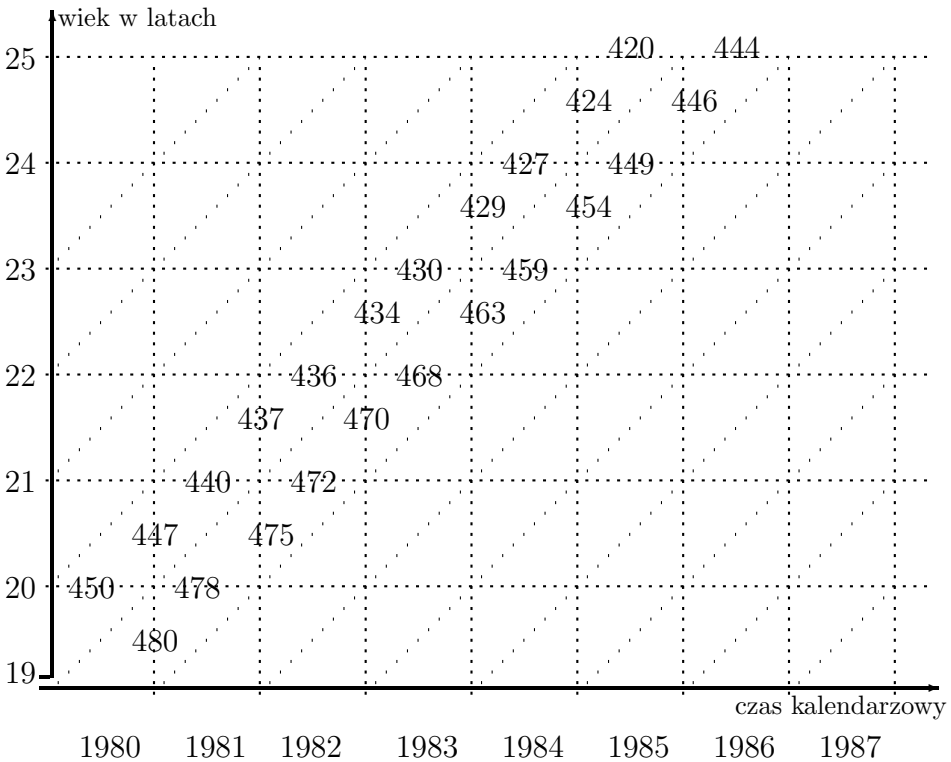
Źródło: GUS

Na podstawie analizy rysunku 22 możemy stwierdzić, że subpopulacje mężczyzn i kobiet w Polsce różnią się udziałami poszczególnych grup wieku (względne niedobory w porównywanych proporcjach zaznaczone zostały symbolicznie liniami przerywanymi). Na przykład, wśród kobiet obserwujemy rosnącą nadwyżkę udziałów grup wieku począwszy od 60 lat w porównaniu do analogicznych udziałów wśród mężczyzn.

Ćwiczenia

Ćwiczenie 2.1

Dany jest fragment siatki demograficznej (rysunek 23), przedstawiający umowne liczby osób (w tys.) z dwóch generacji, dożywających wieku $x = 20, \dots, 25$ lat (wartości na odcinkach poziomych siatki) oraz liczby osób (w tys.), które w dniu 1 stycznia poszczególnych lat były w wieku $x = 19, 20, \dots, 24$ ukończonych lat (wartości na odcinkach pionowych siatki).



Rysunek 23. Fragment siatki demograficznej oraz liczby dożywających I i II rodzaju (dane umowne)

Korzystając z odpowiednich wzorów, obliczyć i zinterpretować:

1. Kohortowe, cząstkowe współczynniki zgonów dla generacji osób urodzonych w 1960 roku (wg rocznych grup wieku 20, 21, ..., 24 lat),
2. Kohortowo-przekrojowe współczynniki zgonów dla generacji urodzonych w 1961 roku, osobno dla lat 1981, ..., 1985 i grup wieku uwidoczonych na rysunku,

3. Prawdopodobieństwa zgonu w ciągu roku dla osób z obu generacjach łącznie, będących w grupie wieku:

- a) 21 ukończonych lat, b) 22 ukończonych lat.

4. Prawdopodobieństwa perspektywiczne zgonu w ciągu roku kalendarzowego dla osób z obu generacji łącznie, które dożyły do 1 stycznia roku:

- a) 1981, b) 1982, c) 1983.

Ćwiczenie 2.2

Na podstawie siatki demograficznej przedstawionej na rysunku 6 (strona 20) obliczyć:

1. Kohortowy współczynnik płodności w grupie wieku 21 ukończonych lat w generacji kobiet urodzonych w 1989 roku,
2. Kohortowy współczynnik zgonów dla generacji kobiet urodzonych w roku 1990 i będących w grupie wieku 23 ukończonych lat,
3. Prawdopodobieństwo zgonu w ciągu roku w grupie wieku 24 ukończonych lat w obu generacjach kobiet.

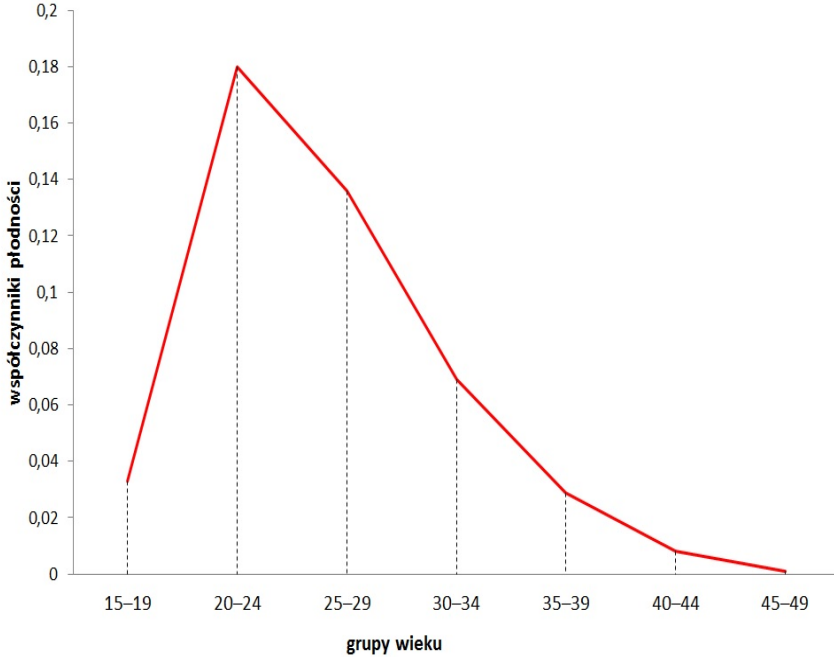
Ćwiczenie 2.3

Tablica 5 przedstawia cząstkowe współczynniki płodności w Polsce w roku 1980, natomiast rysunek 24 ilustruje profil płodności wg wieku w tym roku.

Tablica 5. Cząstkowe współczynniki płodności w Polsce w roku 1980

Grupa wieku	${}_5f_x$
15–19	0,033
20–24	0,180
25–29	0,136
30–34	0,069
35–39	0,029
40–44	0,008
45–49	0,001

Źródło: dane GUS, stat.gov.pl



Rysunek 24. Profil płodności w Polsce w roku 1980

Źródło: dane GUS, obliczenia własne.

1. Obliczyć analogiczne współczynniki cząstkowe dla lat 2010, 2015, korzystając z danych statystycznych zamieszczonych w Banku Danych Lokalnych na stronie GUS (stat.gov.pl) na temat:
 - liczby urodzeń w podziale wg pięcioletnich grup wieku matek (jak w tabelicy 5),
 - liczby kobiet w poszczególnych grupach wieku w połowie roku.
2. Sporządzić wykres profili płodności dla lat 2010 i 2015 oraz ocenić różnice w stosunku do roku 1980.
3. Obliczyć i zinterpretować ogólne współczynniki płodności *GFR*, osobno dla roku 2010 i roku 2015.
4. Dokonać standaryzacji współczynników *GFR* dla lat 2010 i 2015, przyjmując:

- a) strukturę kobiet wg wieku na poziomie średnim z analizowanych dwóch okresów,
- b) średnie współczynniki cząstkowe z obu okresów.

Wyniki porównać z uzyskanymi w punkcie 3.

5. Dokonać dekompozycji różnicy pomiędzy ogólnymi współczynnikami mi płodności, o których mowa w punkcie 3. Wyniki zinterpretować.

Ćwiczenie 2.4

Tablica 6 przedstawia liczby zgonów wg płci zarejestrowane w Polsce w latach 1980 i 2007 oraz odpowiadające im cząstkowe współczynniki zgonów.

Tablica 6. Liczby zgonów oraz cząstkowe współczynniki zgonów wg płci w Polsce w latach 1980 i 2007

Płeć	Zgony		Zgony na 1000 osób	
	1980	2007	1980	2007
kobiety	162 331	174 985	8,89	8,88
mężczyźni	190 833	202 241	11,01	10,98

Źródło: dane GUS, stat.gov.pl

1. Wyznaczyć i zinterpretować ogólne współczynniki zgonów *CDR*, osobno dla roku 1980 i roku 2007.
2. Dokonać standaryzacji współczynników *CDR* dla lat 1980 oraz 2007, zakładając:
 - a) strukturę ludności wg płci na poziomie średnim z analizowanych okresów,
 - b) średnie współczynniki cząstkowe z obu okresów.

Wyniki porównać z uzyskanymi w punkcie 1.

3. Dokonać dekompozycji różnicy pomiędzy ogólnymi współczynnikami mi zgonów, o których mowa w punkcie 1. Wyniki zinterpretować.

Rozdział 3

TABLICE TRWANIA ŻYCIA

3.1. Wprowadzenie

Tablica trwania życia TTŻ (*life table*) jest jednym z najważniejszych narzędzi analitycznych w demografii [Bolesławski 1973]. To model teoretyczny ilustrujący proces wymierania generacji o umownej liczebności początkowej.

TTŻ można także odnieść do generacji rzeczywistej (rzeczywistej kohorty urodzeniowej). Mówimy wówczas o tablicy kohortowej. Jednak konstruowanie takiej tablicy wymaga gromadzenia danych dotyczących umieralności przez cały okres życia danej kohorty, co w przypadku populacji ludzkich może trwać 100 lub więcej lat. Z tego powodu tablice kohortowe stosowane są bardzo rzadko. Częściej konstruuje się tzw. tablice przekrojowe (bieżące), które bazują na danych bieżących, np. na przekrojowych współczynnikach umieralności odnotowanych w rzeczywistej populacji osób żyjących jednocześnie w tym samym okresie kalendarzowym. Przyjmuje się wówczas, że współczynniki te opisują proces wymierania pewnej hipotetycznej kohorty urodzeniowej. Przy założeniu braku migracji, stałej liczby urodzeń oraz niezmiennego porządku wymierania przekrojowa tablica trwania życia może być traktowana jako model tzw. populacji stacjonarnej [zob. paragraf 3.6].

3.2. Podstawowe charakterystyki TTŻ

Tablica trwania życia składa się z kilku kolumn, z których pierwsza reprezentuje grupy wieku, np. roczne, dwuletnie, pięcioletnie. Pozostałe kolumny zawierają wartości parametrów charakteryzujących proces wymierania generacji w podziale na zadane przedziały wieku, m.in.

liczbę osób dożywających do początku danego przedziału, średnie dalsze trwanie życia czy też liczbę zgonów w poszczególnych przedziałach wieku itp.

Rozważmy w pierwszej kolejności tablicę pełną, tj. tablicę dla rocznych przedziałów wieku $[x, x + 1)$. Niech x będzie identyfikatorem poszczególnych przedziałów o wartościach całkowitych $0, 1, 2, \dots$. Zmienną x możemy także interpretować jako wiek w ukończonych latach.

Do najważniejszych charakterystyk tablic trwania życia należą:

1. Liczba l_x dożywających wieku x ukończonych lat, przy czym liczebność początkowa l_0 jest zadana z góry (np. $l_0 = 100\ 000$) oraz mają miejsce relacje

$$l_0 > l_1 > l_2 \dots > l_X = 0, \quad (3.1)$$

gdzie X oznacza górną nieprzekraczalną granicę wieku.

2. Liczba zgonów d_x w wieku x ukończonych lat, tj. w grupie wieku $[x, x + 1)$ lat, będąca różnicą liczb dożywających wieku x oraz $x + 1$ lat, czyli

$$d_x = l_x - l_{x+1}, \quad (3.2)$$

z czego wynika również relacja

$$l_{x+1} = l_x - d_x. \quad (3.3)$$

3. Średni czas ekspozycji na ryzyko zgonu a_x w przedziale wieku $[x, x + 1)$ lat osób zmarłych w tym przedziale wieku, przy czym często przyjmuje się $a_x = \frac{1}{2}$ [zob. sekcja 3.5],

4. Czas ekspozycji K_x na ryzyko zgonu w grupie wieku $[x, x + 1)$, definiowane jako

$$K_x = l_{x+1} + a_x \cdot d_x \quad (3.4)$$

lub równoważnie

$$K_x = l_x - (1 - a_x) \cdot d_x. \quad (3.5)$$

5. Współczynnik zgonów m_x w grupie wieku $[x, x + 1)$ lat

$$m_x = \frac{d_x}{K_x}. \quad (3.6)$$

6. Fundusz T_x osobołat lat życia osób w wieku co najmniej x ukończonych lat

$$T_x = \sum_{u=x}^{\infty} K_u. \quad (3.7)$$

7. Prawdopodobieństwo zgonu q_x w ciągu roku dla osób, które dożyły x -tych urodzin

$$q_x = \frac{d_x}{l_x} \quad (3.8)$$

lub równoważnie

$$q_x = \frac{m_x}{1 + (1 - a_x) \cdot m_x}. \quad (3.9)$$

8. Prawdopodobieństwo przeżycia p_x kolejnego roku dla osób, które dożyły x -tych urodzin

$$p_x = 1 - q_x \quad (3.10)$$

lub

$$p_x = \frac{l_{x+1}}{l_x}. \quad (3.11)$$

9. Średnie dalsze trwanie życia e_x dla osób, które dożyły wieku x lat

$$e_x = \frac{T_x}{l_x}. \quad (3.12)$$

W przypadku przedziałów wieku $[x, x+n)$ dla $n > 1$ ($n \in \mathbf{N}$), konstruowane są tzw. tablice skrócone. Zmienna x przyjmuje wówczas wartości całkowite $0, n, 2n, \dots$. Modyfikacji ulegają także oznaczenia i formuły obliczania niektórych parametrów tablicowych, wymienione poniżej.

Ad 2. Liczba zgonów ${}_n d_x$ w grupie wieku $[x, x+n)$ lat, będąca różnicą liczb dożywających wieku x oraz $x+n$ lat

$${}_n d_x = l_x - l_{x+n}. \quad (3.13)$$

Ad 3. Średnia liczba osobołat życia ${}_n a_x$ w przedziale $[x, x+n)$ lat dla osób zmarłych w tym przedziale wieku – na ogół przybliżamy ją jako ${}_n a_x = \frac{n}{2}$.

Ad. 4. Czas ekspozycji ${}_nK_x$ na ryzyko zgonu w grupie wieku $[x, x+n)$

$${}_nK_x = n \cdot l_{x+n} + {}_n a_x \cdot {}_n d_x \quad (3.14)$$

lub równoważnie

$${}_nK_x = n \cdot l_x - (n - {}_n a_x) \cdot {}_n d_x. \quad (3.15)$$

Ad 5. Współczynnik zgonów ${}_n m_x$ w grupie wieku $[x, x+n)$ lat

$${}_n m_x = \frac{{}_n d_x}{{}_n K_x}. \quad (3.16)$$

Ad 7. Prawdopodobieństwo ${}_n q_x$ zgonu w ciągu n kolejnych lat dla osób, które dożyły wieku x lat

$${}_n q_x = \frac{{}_n d_x}{l_x} \quad (3.17)$$

lub

$${}_n q_x = \frac{n \cdot {}_n m_x}{1 + (n - {}_n a_x) \cdot {}_n m_x}. \quad (3.18)$$

Ad 8. Prawdopodobieństwo ${}_n p_x$ przeżycia $x+n$ -tych urodzin dla osób, które dożyły wieku x lat

$${}_n p_x = 1 - {}_n q_x \quad \text{lub} \quad {}_n p_x = \frac{l_{x+n}}{l_x}. \quad (3.19)$$

Warto zauważyć, że parametry l_x, T_x, e_x skróconej TTŻ obliczone dla $x = 0, n, 2n, \dots$ są zgodne co do wartości z analogicznymi parametrami tablicy pełnej.

3.3. Tablica kohortowa

Kohortowa tablica trwania życia przedstawia proces wymierania **rzeczywistej kohorty urodzeniowej**, którą jest zazwyczaj kohorta roczna (generacja). Przyjmijmy dla uproszczenia, że na zmiany stanu liczebnego badanej kohorty nie mają wpływu ruchy wędrowkowe (migracje). Załóżmy również, że posiadamy dane pochodzące z obserwacji kohorty rzeczywistej, dotyczące:

- 1) **liczb zgonów I rodzaju** ${}_nD_x^{koh.s}$ w grupach wieku $[x, x+n)$ lat w danej kohorcie, dalej oznaczonych w skrócie jako ${}_nD_x$,
- 2) **liczb dożywających I rodzaju** $U_x^{koh.s}$ w danej kohorcie, w skrócie oznaczonych dalej jako U_x ,
- 3) **współczynników** ${}_na_x$, reprezentujących średni czas ekspozycji na ryzyko zgonu w grupach wieku $[x, x+n)$ lat dla tych osób z kohorty rzeczywistej, które zmarły w rozważanej grupie wieku.

Liczba zgonów w przypadku rocznych grup wieku może być niekiedy relatywnie mała, dlatego wskazane jest w takich przypadkach korzystanie z szerszych przedziałów wieku $[x, x+n)$, np. dla $n=3$. Należy także zaznaczyć, że liczba zgonów ${}_nD_x$ w rzeczywistej kohorcie nie jest zwykle równa liczbie zgonów ${}_nd_x$ w tablicy trwania życia, która opisuje proces wymierania kohorty o umownej liczebności początkowej l_0 . Również wartości parametrów l_x nie pokrywają się z wartościami U_x liczb dożywających I rodzaju w kohorcie rzeczywistej. Niemniej jednak, parametry ${}_nd_x$ oraz l_x kohortowej tablicy trwania życia pozostają w ustalonej relacji odpowiednio do liczb ${}_nD_x$ oraz U_x . Relację tę można wyrazić wzorami

$$l_x = \frac{l_0}{U_0} \cdot U_x, \quad {}_nd_x = \frac{l_0}{U_0} \cdot {}_nD_x. \quad (3.20)$$

Zgodnie z podanymi wyżej założeniami 1–3 załóżmy dalej, że znamy wartości ${}_nD_x$, U_x oraz ${}_na_x$ w grupach wieku $[x, x+n)$ w kohorcie rzeczywistej. Wówczas cząstkowe współczynniki zgonów ${}_nm_x$ wyznaczamy przez analogię do (2.34)

$${}_nm_x = \frac{{}_nD_x}{{}_nK_x} = \frac{{}_nD_x}{n \cdot U_{x+n} + {}_na_x \cdot {}_nD_x}. \quad (3.21)$$

W przypadku braku danych na temat współczynników ${}_na_x$, przyjmujemy przybliżenie ${}_na_x \approx n/2$. Mianownik formuły (3.21) redukuje się wtedy do wyrażenia $\frac{n}{2}(U_x + U_{x+n})$, a cząstkowe współczynniki zgonów wyrażają się przybliżonym wzorem

$${}_nm_x \approx \frac{{}_nD_x}{\frac{n}{2}(U_x + U_{x+n})}. \quad (3.22)$$

Zauważymy, że w mianowniku (3.22) występuje średnia liczba dożywających do początku i końca przedziału wieku $[x, x+n)$, pomnożona przez jego długość n [por. Uwaga 3 w sekcji 2.3].

Podobnie prawdopodobieństwa zgonów ${}_nq_x$ w kohorcie rzeczywistej możemy wyznaczyć na podstawie danych o liczbie zgonów ${}_nD_x$ oraz liczbie U_x dożywających I rodzaju, tj. korzystając ze wzorów

$${}_nq_x = \frac{{}_nD_x}{U_x} \quad (3.23)$$

lub

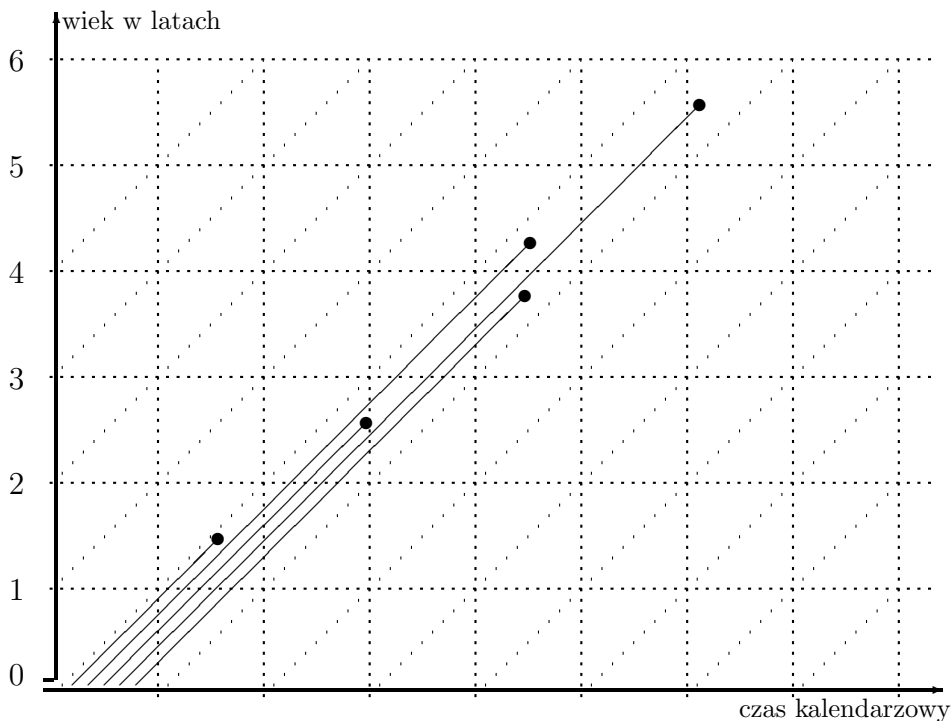
$${}_nq_x = 1 - \frac{U_{x+n}}{U_x}. \quad (3.24)$$

W praktyce, kohortowe współczynniki zgonów ${}_nm_x$ czy też prawdopodobieństwa zgonów ${}_nq_x$ poddawane są często dodatkowym procedurom wygładzania, w celu zredukowania zakłóceń losowych. Zakłada się bowiem, że wielkości te nie zmieniają się w sposób skokowy wraz z przechodzeniem do kolejnych grup wieku, ale raczej w sposób ciągły. Stosowane jest np. wygładzanie metodą ważonych średnich ruchomych. Następnie, na podstawie wygładzonych parametrów ${}_nm_x$ czy ${}_nq_x$ pozostałe wskaźniki kohortowej tablicy trwania życia ${}_nd_x$, l_x , ${}_nK_x$, T_x , e_x wyznaczane są zgodnie z formułami przedstawionymi w sekcji 3.2. Jednak zależnie od stopnia wygładzenia parametrów wyjściowych, relacje (3.20) mogą stać się mniej lub bardziej przybliżone.

Warto dodać, że średnie trwanie życia noworodka e_0 uzyskane na podstawie danych z kohorty rzeczywistej określa również średni wiek \bar{x} w momencie śmierci poszczególnych jednostek należących do danej kohorty. W przypadku ujęcia przekrojowego równość taka nie zachodzi. Wyjątkiem jest przypadek tzw. ludności stacjonarnej [rozdział 6].

Przykład 11. Zbudujemy tablicę trwania życia na podstawie linii życia przedstawionych na rysunku 25. Zakładamy przy tym, że mamy do czynienia z kohortą o dużej liczebności. Przyjmujemy zatem dla uproszczenia, że każda linia życia na rysunku 25 reprezentuje np. 1 mln osób.

Obliczenia rozpoczynamy od wyznaczenia prawdopodobieństw zgonów (3.23). Następnie, wykorzystamy te prawdopodobieństwa do znalezienia pozostałych parametrów kohortowej TTŻ dla rocznych i dwuletnich grup wieku, przyjmując umowną liczebność początkową $l_0 = 100\ 000$. Wyniki obliczeń zawarte są w tablicach 7 (wersja pełna) i 8 (wersja skrócona). Zauważymy, że parametry l_x , T_x oraz e_x w tablicy skróconej równe są analogicznym parametrom z tablicy pełnej dla $x = 0, 2, 4$.



Rysunek 25. Fragment siatki demograficznej i przykładowe linie życia

Tablica 7. Pełna TTŻ dla rocznych grup wieku

x	q_x	l_x	d_x	K_x	a_x	T_x	e_x
0	0	100 000	0	100 000	0,0	354 000	3,54
1	1/5	100 000	20 000	90 000	0,5	254 000	2,54
2	1/4	80 000	20 000	72 000	0,6	164 000	2,05
3	1/3	60 000	20 000	56 000	0,8	92 000	1,53
4	1/2	40 000	20 000	24 000	0,2	36 000	0,90
5	1	20 000	20 000	12 000	0,6	12 000	0,60

Źródło: obliczenia na podstawie danych z rysunku 25

Tablica 8. Skrócona TTŻ dla dwuletnich grup wieku

x	${}_2q_x$	l_x	${}_2d_x$	${}_2K_x$	${}_2a_x$	T_x	e_x
0	1/5	100 000	20 000	190 000	1,5	354 000	3,54
2	1/2	80 000	40 000	128 000	1,2	164 000	2,05
4	1	40 000	4 0000	36 000	0,9	36 000	0,90

Źródło: obliczenia na podstawie danych z rysunku 25

Łatwo sprawdzić, że oczekiwany czas życia e_0 jest w tym przypadku równy średniemu wiekowi \bar{x} w chwili zgonu. Mamy bowiem

$$\bar{x} = \frac{1,5 + 2,6 + 3,8 + 4,2 + 5,6}{5} = 3,54 = e_0.$$

3.4. Tablica przekrojowa

Przekrojowa TTŻ ilustruje proces wymierania hipotetycznej generacji o ustalonej liczebności l_0 . Zakładamy, że w generacji tej intensywność zgonów w poszczególnych przedziałach wieku odzwierciedla intensywność zgonów odnotowaną we wszystkich rzeczywistych generacjach żyjących w roku t . Innymi słowy, przekrojowa TTŻ jest konstruowana na podstawie danych pochodzących z rejestracji liczby zgonów i danych o liczbie żyjących w ustalonym roku.

Przekrojowe TTŻ dla Polski publikowane są przez Główny Urząd Statystyczny, osobno dla mężczyzn i kobiet. Metoda konstrukcji przekrojowych TTŻ stosowana przez GUS uwzględnia m.in. korektę liczby ludności spowodowaną migracjami oraz zakłada wygładzanie prawdopodobieństw zgonów wyznaczonych z danych empirycznych za pomocą średnich ruchomych. Poniżej przedstawimy etapy konstrukcji tablic dla rocznych grup wieku wg zasad GUS [*Tablice trwania życia 2017*, s. 8–12].

W pierwszym kroku wyznaczane są prawdopodobieństwa zgonów q_x dla $x = 0, 1, \dots, X$. Wymaga to przedstawienia przekrojowych danych o liczbie zgonów w układzie przekrojowo-kohortowym.

Przyjmijmy następujące oznaczenia:

- $L_x(\tau)$ – liczba osób w wieku x ukończonych lat w momencie τ ,
- $B(t)$ – liczba urodzeń żywych w roku t ,
- $D_x^{ml}(t)$ – liczba zgonów w roku t w wieku x ukończonych lat wśród urodzonych w roku $t - x - 1$ (zgoni młodsi),
- $D_x^{st}(t)$ – liczba zgonów w roku t w wieku x ukończonych lat wśród urodzonych w roku $t - x$ (zgoni starsi),
- $R_x(t)$ – korekta liczby ludności spowodowana migracjami w roku t osób urodzonych w roku $t - x$,
- $R_{x+1}(t)$ – korekta liczby ludności spowodowana migracjami w roku t wśród osób urodzonych w roku $t - x - 1$.

Rozważmy dwa sąsiednie lata kalendarzowe oznaczone umownie jako $t - 1 = [\tau_0, \tau_1]$ oraz $t = [\tau_1, \tau_2]$. Symbol τ_0 oznacza początek roku $t - 1$, τ_1 – koniec roku $t - 1$ i jednocześnie początek roku t , natomiast τ_2 – koniec roku t .

Skorzystamy dalej z relacji wyrażonych w postaci równań bilansowych

$$L_x(\tau_2) = \begin{cases} B(t) - D_x^{st}(t) - R_x(t), & x = 0, \\ L_{x-1}(\tau_1) - D_{x-1}^{ml}(t) - D_x^{st}(t) - R_x(t), & x \geq 1. \end{cases} \quad (3.25)$$

Stąd otrzymujemy

$$R_x(t) = \begin{cases} B(t) - L_x(\tau_2) - D_x^{st}(t), & x = 0, \\ L_{x-1}(\tau_1) - L_x(\tau_2) - D_{x-1}^{ml}(t) - D_x^{st}(t), & x \geq 1. \end{cases} \quad (3.26)$$

Prawdopodobieństwa q_x dla $x = 0, 1, \dots, 84$ obliczamy, korzystając z formuły

$$q_x = 1 - (1 - q'_x)(1 - q''_x), \quad (3.27)$$

gdzie q'_x oraz q''_x są zdefiniowane wzorami

$$q'_x = \frac{D_x^{ml}(t) + D_x^{ml}(t-1)}{L_x(\tau_1) - \frac{1}{2}R_{x+1}(t) + L_x(\tau_0) - \frac{1}{2}R_{x+1}(t-1)}, \quad (3.28)$$

$$q''_x = \frac{D_x^{st}(t) + D_x^{st}(t-1)}{L_x(\tau_2) + D_x^{st}(t) + \frac{1}{2}R_x(t) + L_x(\tau_1) + D_x^{st}(t-1) + \frac{1}{2}R_x(t-1)}.$$

Prawdopodobieństwa q_x dla najstarszych grup wieku $x = 85, 86, \dots$ są wyznaczane poprzez dopasowanie funkcji wielomianowo-wykładniczej

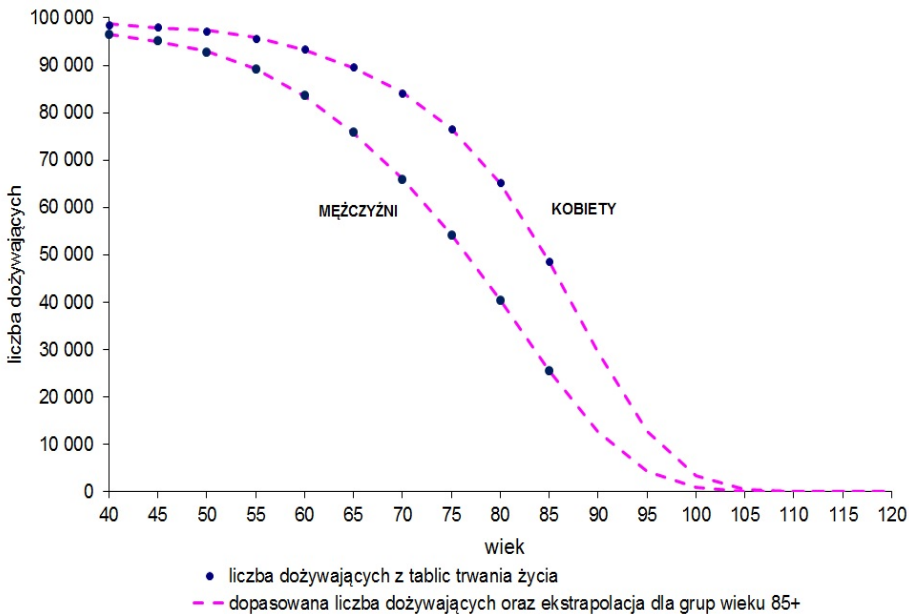
$$f(x) = l_0 \exp\{-b_0 - b_1x - b_2x^2 - \dots - b_5x^5\} \quad (3.29)$$

do liczb l_x dożywających wieku $x = 40, 45, 50, \dots, 85$ lat, dla zadanej liczebności początkowej $l_0 = 100\ 000$. Wykorzystywana jest do tego celu metoda najmniejszych kwadratów i algorytm optymalizacji nieliniowej Marquardta.

W kolejnym kroku dokonujemy ekstrapolacji funkcji (3.29), uzyskując liczby dożywających $\tilde{l}_x = f(x)$ dla $x > 85$.

Przykład 12. Rysunek 26 przedstawia funkcję (3.29) dopasowaną metodą optymalizacji nieliniowej Marquardta do liczb dożywających l_x zaczerpniętych z tablic trwania życia dla Polski dla $x = 40, 45, \dots, 85$ lat (dane GUS, 2017 rok), a także ekstrapolację tej funkcji dla roczników od $x = 85$ lat do $x = 120$ lat, osobno dla mężczyzn i kobiet.

Jak można wywnioskować na podstawie układu krzywych na rysunku 26, dopasowanie funkcji wielomianowo-wykładniczej do liczb dożywających l_x dla $x = 40, 45, \dots, 85$ jest stosunkowo dokładne. Potwierdza to także m.in. średnie względne odchylenie wartości tablicowych l_x od wartości dopasowanych, które kształtuje się w tym przypadku na poziomie ok. 0,001 dla mężczyzn i ok. 0,002 dla kobiet.



Rysunek 26. Dopasowanie funkcji wielomianowo-wykładniczej do liczb dożywających zaczerpniętych z TTŻ dla Polski z 2017 roku

Źródło: obliczenia własne na podstawie TTŻ 2017, GUS

Na podstawie \tilde{l}_x wyznaczamy także prawdopodobieństwa zgonów dla $x = 85, 86, \dots, 120$, korzystając z formuły

$$q_x = \frac{\tilde{l}_x - \tilde{l}_{x+1}}{\tilde{l}_x}. \quad (3.30)$$

W celu wygładzenia prawdopodobieństw zgonu q_x przeprowadza się dodatkowo zabieg wyrównywania za pomocą tzw. ważonych średnich ruchomych (z wyjątkiem q_0). Stosowane jest w tym celu ruchome

dopasowanie paraboliczne z różną liczbą wyrazów, co sprowadza się do następujących średnich ruchomych ważonych:

- w wieku $x = 1$ roku – średnia z wartości od 1 do 5 lat o wagach: 0,88571; 0,25714; -0,08571; -0,14286; 0,08571,
- w wieku $x = 2$ lat – średnia z wartości od 1 do 5 lat o wagach: 0,25714; 0,37143; 0,34286; 0,17143; -0,14286,
- w wieku $x = 3$ lat – scentrowana pięciowyrazowa średnia o wagach: -0,08571; 0,34286; 0,48571; 0,34286; -0,08571,
- w wieku od $x = 4$ do $x = 29$ lat – scentrowana siedmiowyrazowa średnia o wagach: -0,09524; 0,14286; 0,28571; 0,33333; 0,28571; 0,14286; -0,09524,
- w wieku od $x = 30$ do $x = 89$ lat – scentrowana dziewięciowyrazowa średnia o wagach: -0,09091; 0,06061; 0,16883; 0,23377; 0,25541; 0,23377; 0,16883; 0,06061; -0,09091.

Cykl wygładzania prawdopodobieństw q_x powtarza się trzykrotnie, co prowadzi do otrzymania prawdopodobieństw wygładzonych.

Etapy konstrukcji przekrojowej TTŻ

1. Wyznaczamy prawdopodobieństwa przeżycia p_x na podstawie wygładzonych prawdopodobieństw q_x , wg wzoru

$$p_x = 1 - q_x. \quad (3.31)$$

2. Określamy liczby l_{x+1} dożywających do wieku $x + 1$ lat

$$l_{x+1} = l_x \cdot p_x, \quad (3.32)$$

przyjmując początkową liczebność l_0 (np. $l_0 = 100\,000$).

3. Obliczamy liczby zgonów d_x w grupie osób w wieku x ukończonych lat

$$d_x = l_x \cdot q_x = l_x - l_{x+1}. \quad (3.33)$$

4. Ustalamy liczbę osobołat życia K_x w grupie wieku $[x, x + 1)$ dla $x \geq 1$ (tzw. ludność stacjonarną)

$$K_x = l_{x+1} + a_x \cdot d_x, \quad (3.34)$$

lub równoważnie

$$K_x = l_x - (1 - a_x) \cdot d_x. \quad (3.35)$$

Wartości współczynników a_x zastępujemy przybliżeniem $a_x \approx \frac{1}{2}$. Mamy wtedy

$$K_x = l_{x+1} + \frac{1}{2}d_x \quad (3.36)$$

lub równoważnie

$$K_x = l_x - \frac{1}{2}d_x. \quad (3.37)$$

Dla $x = 0$ przyjmujemy

$$K_0 = l_0 \cdot (1 - q_0''). \quad (3.38)$$

5. Znajdujemy fundusz osobołat lat T_x osób w wieku x (tzw. ludność stacjonarną skumulowaną)

$$T_x = \sum_{u \geq x} K_u. \quad (3.39)$$

6. Wyznaczamy średnie dalsze trwanie życia e_x

$$e_x = \frac{T_x}{l_x}. \quad (3.40)$$

Analogiczny sposób postępowania można zastosować także przy budowie TTŻ w podziale na szersze przedziały wieku $[x, x + n)$ lat, czyli tzw. skróconych przekrojowych tablic trwania życia.

Wyjściowe prawdopodobieństwa zgonu ${}_nq_x$ są obliczane w tym przypadku na podstawie tablicy pełnej, korzystając z relacji

$${}_nq_x = 1 - \frac{l_{x+n}}{l_x}. \quad (3.41)$$

Następnie wyznaczane są pozostałe parametry tablicowe.

Ze względu na różnice w natężeniu umieralności kobiet i mężczyzn, tablice trwania życia tworzone są zwykle osobno dla obu płci. Fragmenty tablic przekrojowych GUS dla 2017 roku dla mężczyzn i kobiet przedstawiają tablice 9 i 10.

Tablica 9. Fragment przekrojowej tablicy trwania życia
mężczyźni (2017 rok)

x	l_x	q_x	d_x	K_x	T_x	e_x
0	100 000	0,00446	446	99 611	7 396 267	73,96
1	99 554	0,00028	28	99 540	7 296 656	73,29
2	99 526	0,00019	19	99 516	7 197 117	72,31
3	99 506	0,00013	13	99 500	7 097 601	71,33
4	99 493	0,00010	10	99 488	6 998 101	70,34
5	99 482	0,00009	9	99 478	6 898 614	69,35
6	99 473	0,00009	9	99 469	6 799 136	68,35
7	99 464	0,00009	9	99 460	6 699 667	67,36
8	99 455	0,00009	9	99 450	6 600 207	66,36
9	99 446	0,00009	9	99 441	6 500 757	65,37
10	99 436	0,00010	10	99 432	6 401 316	64,38
11	99 427	0,00010	10	99 422	6 301 884	63,38
12	99 417	0,00011	11	99 411	6 202 462	62,39
13	99 405	0,00013	13	99 399	6 103 051	61,40
14	99 392	0,00017	16	99 384	6 003 653	60,40
15	99 376	0,00022	22	99 365	5 904 268	59,41
16	99 354	0,00032	32	99 338	5 804 904	58,43
17	99 322	0,00045	45	99 299	5 705 566	57,45
18	99 277	0,00059	58	99 248	5 606 267	56,47
19	99 219	0,00071	71	99 183	5 507 019	55,50
20	99 148	0,00080	79	99 108	5 407 835	54,54
21	99 069	0,00085	84	99 027	5 308 727	53,59
22	98 985	0,00088	87	98 941	5 209 700	52,63
23	98 898	0,00090	89	98 853	5 110 759	51,68
24	98 809	0,00093	92	98 763	5 011 906	50,72
25	98 716	0,00097	96	98 668	4 913 143	49,77
26	98 621	0,00101	100	98 571	4 814 475	48,82
27	98 521	0,00105	103	98 469	4 715 904	47,87
28	98 417	0,00109	107	98 364	4 617 435	46,92
29	98 310	0,00114	113	98 254	4 519 071	45,97
30	98 197	0,00121	119	98 138	4 420 817	45,02
31	98 079	0,00128	126	98 016	4 322 679	44,07
32	97 953	0,00136	133	97 887	4 224 663	43,13
33	97 820	0,00145	142	97 749	4 126 777	42,19
34	97 679	0,00154	151	97 603	4 029 027	41,25
35	97 528	0,00165	161	97 447	3 931 424	40,31
36	97 367	0,00177	172	97 281	3 833 977	39,38
37	97 195	0,00191	185	97 102	3 736 696	38,45
38	97 009	0,00207	200	96 909	3 639 594	37,52
39	96 809	0,00225	218	96 700	3 542 684	36,59
40	96 591	0,00246	238	96 473	3 445 984	35,68
41	96 354	0,00270	261	96 223	3 349 512	34,76
42	96 093	0,00298	286	95 950	3 253 288	33,86
43	95 807	0,00329	315	95 649	3 157 338	32,96
44	95 491	0,00364	347	95 318	3 061 689	32,06
45	95 144	0,00402	382	94 953	2 966 372	31,18
46	94 761	0,00444	421	94 551	2 871 419	30,30
47	94 341	0,00490	462	94 109	2 776 868	29,43
48	93 878	0,00541	508	93 624	2 682 758	28,58
49	93 371	0,00596	557	93 092	2 589 134	27,73
50	92 814	0,00657	610	92 509	2 496 041	26,89
...
100	941	0,32054	302	790	2191	2,33

Źródło: dane GUS, stat.gov.pl

Tablica 10. Fragment przekrojowej tablicy trwania życia
kobiety (2017 rok)

x	l_x	q_x	d_x	K_x	T_x	e_x
0	100 000	0,00361	361	99 685	8 181 635	81,82
1	99 639	0,00023	23	99 627	8 081 950	81,11
2	99 616	0,00015	15	99 608	7 982 322	80,13
3	99 601	0,00010	10	99 596	7 882 714	79,14
4	99 591	0,00008	8	99 587	7 783 118	78,15
5	99 583	0,00008	8	99 579	7 683 531	77,16
6	99 575	0,00008	8	99 571	7 583 952	76,16
7	99 567	0,00008	7	99 563	7 484 381	75,17
8	99 559	0,00007	7	99 556	7 384 818	74,17
9	99 553	0,00007	7	99 549	7 285 262	73,18
10	99 546	0,00007	7	99 542	7 185 713	72,18
11	99 539	0,00008	8	99 535	7 086 170	71,19
12	99 531	0,00010	10	99 526	6 986 636	70,20
13	99 521	0,00012	12	99 515	6 887 110	69,20
14	99 509	0,00014	14	99 502	6 787 595	68,21
15	99 495	0,00016	16	99 487	6 688 093	67,22
16	99 479	0,00019	19	99 469	6 588 606	66,23
17	99 460	0,00022	22	99 449	6 489 137	65,24
18	99 438	0,00024	24	99 426	6 389 688	64,26
19	99 414	0,00026	25	99 401	6 290 262	63,27
20	99 389	0,00026	25	99 376	6 190 861	62,29
21	99 363	0,00025	25	99 351	6 091 485	61,31
22	99 339	0,00024	24	99 327	5 992 134	60,32
23	99 315	0,00023	23	99 304	5 892 808	59,33
24	99 292	0,00023	23	99 281	5 793 504	58,35
25	99 269	0,00023	23	99 258	5 694 223	57,36
26	99 246	0,00025	24	99 234	5 594 965	56,37
27	99 222	0,00026	26	99 209	5 495 731	55,39
28	99 196	0,00028	28	99 182	5 396 523	54,40
29	99 168	0,00031	31	99 152	5 297 341	53,42
30	99 137	0,00033	33	99 121	5 198 189	52,43
31	99 104	0,00035	35	99 087	5 099 068	51,45
32	99 069	0,00038	38	99 050	4 999 981	50,47
33	99 031	0,00041	41	99 011	4 900 930	49,49
34	98 990	0,00045	45	98 968	4 801 919	48,51
35	98 946	0,00050	49	98 921	4 702 951	47,53
36	98 897	0,00055	54	98 870	4 604 030	46,55
37	98 843	0,00060	60	98 813	4 505 160	45,58
38	98 783	0,00067	66	98 750	4 406 348	44,61
39	98 717	0,00075	74	98 680	4 307 598	43,64
40	98 643	0,00084	83	98 601	4 208 918	42,67
41	98 560	0,00094	92	98 514	4 110 317	41,70
42	98 467	0,00105	103	98 416	4 011 803	40,74
43	98 364	0,00118	116	98 306	3 913 388	39,78
44	98 248	0,00131	129	98 184	3 815 082	38,83
45	98 119	0,00146	144	98 047	3 716 898	37,88
46	97 976	0,00163	160	97 896	3 618 850	36,94
47	97 816	0,00181	177	97 728	3 520 955	36,00
48	97 639	0,00200	196	97 542	3 423 227	35,06
49	97 444	0,00222	216	97 336	3 325 685	34,13
50	97 227	0,00246	239	97 108	3 228 350	33,20
...
100	2 868	0,30613	878	2 429	6902	2,41

Źródło: dane GUS, stat.gov.pl

3.5. Strategie wyboru współczynników ${}_n a_x$

A. Bezpośrednia obserwacja

Sposób obliczania ${}_n a_x$, ${}_n m_x$ lub ${}_n q_x$ poprzez bezpośrednią obserwację jest możliwy do zastosowania w przypadku, gdy mamy dokładne dane o wieku w chwili zgonu badanych osób. Z tego powodu rzadko się go wykorzystuje.

Przykład 13. Załóżmy, że zarejestrowano czas życia w pewnej generacji 5 osób (dane w latach):

$$70, 44; 70, 81; 33, 17; 70, 85; 71, 62.$$

W celu wyznaczenia współczynnika a_x , np. dla grupy wieku $[70, 71)$ lat, wystarczy obliczyć średnią wartość z tych obserwacji, które należą do przedziału $[70, 71)$ lat i odjąć od niej początek przedziału, co daje

$$a_{70} = \frac{1}{3}(70, 44 + 70, 81 + 70, 85) - 70 = 0, 7.$$

Czas ekspozycji na ryzyko zgonu w przedziale $[70, 71)$ lat wynosi

$$K_{70} = 1 + 0, 44 + 0, 81 + 0, 85 = 3, 1.$$

Stąd współczynnik zgonów dla tej grupy wieku równy jest

$$m_{70} = \frac{3}{3, 1} \approx 0, 968,$$

natomiast prawdopodobieństwo zgonu wynosi

$$q_{70} = \frac{3}{4} = 0, 75.$$

Łatwo sprawdzić, że każdą z wielkości a_{70} , m_{70} , q_{70} można wyznaczyć na podstawie pozostałych dwóch

$$a_{70} = 1 + \frac{1}{m_{70}} - \frac{1}{q_{70}} = 1 + \frac{1}{0, 968} - \frac{1}{0, 75} \approx 1 + 1, 033 - 1, 333 = 0, 7.$$

Podobnie

$$m_{70} = \frac{q_{70}}{1 - (1 - a_{70}) \cdot q_{70}} = \frac{0, 75}{1 - (1 - 0, 7) \cdot 0, 75} \approx 0, 968$$

oraz

$$q_{70} = \frac{m_{70}}{1 + (1 - a_{70}) \cdot m_{70}} = \frac{0,968}{1 + (1 - 0,7) \cdot 0,968} \approx 0,75.$$

B. Modele interpolacyjne

Niech l_{x+y} dla $y \in [0, n]$ oznacza liczbę dożywających wieku $x+y$ lat. Liczbę l_{x+y} traktować będziemy jako ciągłą funkcję zmiennej $y \in [0, n]$. Zakładać będziemy dalej liniową bądź wykładniczą postać tej funkcji.

3.5.1. Model interpolacji liniowej

Założmy, że l_{x+y} dla ustalonego x jest liniową funkcją zmiennej y , czyli

$$l_{x+y} = \alpha + \beta y \quad \text{dla } y \in [0, n]. \quad (3.42)$$

Parametry α, β określamy w taki sposób, aby funkcja l_{x+y} przyjmowała zadaną wartość l_x dla $y = 0$ oraz l_{x+n} dla $y = n$. Warunki te możemy zapisać następująco

$$l_x = \alpha \quad \text{dla } y = 0, \quad l_{x+n} = \alpha + \beta n \quad \text{dla } y = n. \quad (3.43)$$

Z warunków tych wynika, że

$$\alpha = l_x, \quad (3.44)$$

$$\beta = \frac{l_{x+n} - l_x}{n} = -\frac{{}_n d_x}{n},$$

gdzie ${}_n d_x$ jest liczbą zgonów w przedziale wieku $[x, x+n)$.

W rezultacie (3.42) możemy zapisać wzorem

$$l_{x+y} = \alpha + \beta y = l_x - \frac{{}_n d_x}{n} y. \quad (3.45)$$

Obliczymy teraz czas ekspozycji ${}_n K_x$ na ryzyko zgonu w przedziale $[x, x+n)$. Gdy l_{x+n} jest funkcją całkowalną, wówczas czas ekspozycji ${}_n K_x$ obliczamy jako całkę oznaczoną z tej funkcji na przedziale $[0, n]$.

W tym przypadku funkcja l_{x+y} ma postać liniową (3.42), a zatem jest funkcją całkowalną. Wobec tego mamy

$$\begin{aligned} {}_nK_x &= \int_0^n l_{x+y} dy = \int_0^n \left(l_x - \frac{{}_n d_x}{n} y \right) dy = \\ &= n l_x - \frac{{}_n d_x}{n} \frac{1}{2} y^2 \Big|_0^n = n \cdot l_x - \frac{n}{2} \cdot {}_n d_x. \end{aligned} \quad (3.46)$$

Ponieważ l_x, l_{x+n} łączy relacja

$$l_x = l_{x+n} + {}_n d_x, \quad (3.47)$$

więc otrzymujemy także

$${}_n K_x = n \cdot l_{x+n} + \frac{n}{2} \cdot {}_n d_x. \quad (3.48)$$

Z porównania otrzymanego wyniku z prawą stroną wzoru (3.15) definiującego czas ekspozycji ${}_n K_x$, tj. z wyrażeniem $n \cdot l_{x+n} + {}_n a_x \cdot {}_n d_x$ wnioskujemy ostatecznie, że współczynnik ${}_n a_x$ równy jest

$${}_n a_x = \frac{n}{2}. \quad (3.49)$$

Tym samym w modelu interpolacji liniowej formuła (2.46), definiująca relację pomiędzy prawdopodobieństwem ${}_n q_x$ oraz cząstkowym współczynnikiem zgonów ${}_n m_x$, sprowadza się do postaci

$${}_n q_x = \frac{n \cdot {}_n m_x}{1 + (n - \frac{n}{2}) \cdot {}_n m_x} = \frac{2 \cdot n \cdot {}_n m_x}{2 + n \cdot {}_n m_x}. \quad (3.50)$$

W szczególnym przypadku dla $n = 1$, otrzymujemy

$$q_x = \frac{m_x}{1 + \frac{1}{2} m_x} = \frac{2 \cdot m_x}{2 + m_x}. \quad (3.51)$$

3.5.2. Model interpolacji wykładniczej

Przyjmijmy, że l_{x+y} jest funkcją wykładniczą zmiennej $y \in [0, n]$, określoną wzorem

$$l_{x+y} = \alpha \beta^y, \quad y \in [0, n], \quad (3.52)$$

przy warunkach ograniczających

$$l_x = \alpha \quad \text{dla } y = 0, \quad (3.53)$$

$$l_{x+n} = \alpha\beta^n \quad \text{dla } y = n,$$

przy czym l_x oraz l_{x+n} są zadane z góry.

Z powyższych warunków wynika, że

$$\begin{aligned} \alpha &= l_x, \\ \beta &= \left(\frac{l_{x+n}}{l_x} \right)^{\frac{1}{n}}. \end{aligned} \quad (3.54)$$

Stąd funkcja l_{x+y} zdefiniowana w (3.52) ma postać

$$l_{x+y} = l_x \left(\frac{l_{x+n}}{l_x} \right)^{\frac{y}{n}}. \quad (3.55)$$

Oznaczmy

$$1 - {}_nq_x = {}_np_x. \quad (3.56)$$

Z faktu, że

$$\frac{l_{x+n}}{l_x} = {}_np_x, \quad (3.57)$$

otrzymujemy na podstawie (3.55)

$$l_{x+y} = l_x ({}_np_x)^{\frac{y}{n}}. \quad (3.58)$$

Obliczymy czas ekspozycji ${}_nK_x$ jako całkę na przedziale $[0, n]$ z funkcji l_{x+y} . Mamy

$$\begin{aligned} {}_nK_x &= \int_0^n l_{x+y} dy = \int_0^n l_x ({}_np_x)^{\frac{y}{n}} dy = \\ &= l_x \int_0^n \exp\left\{ \frac{y}{n} \ln {}_np_x \right\} dy. \end{aligned} \quad (3.59)$$

Ostatnie przekształcenie pod całką wynika z faktu, że

$$\alpha^z \equiv e^{z \ln \alpha} \quad \text{dla dowolnej, dodatniej stałej } \alpha. \quad (3.60)$$

Dokonyamy zamiany zmiennych. Niech

$$z = \frac{y}{n} \ln {}_n p_x, \quad \text{stad} \quad dy = \frac{n}{\ln {}_n p_x} dz. \quad (3.61)$$

Mamy wtedy

$$\begin{aligned} {}_n K_x &= n \cdot l_x \int_0^{\ln {}_n p_x} \frac{e^z}{\ln {}_n p_x} dz = \\ &= \frac{n \cdot l_x}{\ln {}_n p_x} \int_0^{\ln {}_n p_x} e^z dz = \frac{n \cdot l_x}{\ln {}_n p_x} e^z \Big|_0^{\ln {}_n p_x} = \\ &= \frac{n \cdot l_x}{\ln {}_n p_x} (e^{\ln {}_n p_x} - e^0) = \frac{n \cdot l_x}{\ln {}_n p_x} ({}_n p_x - 1) = \\ &= -\frac{n \cdot l_x}{\ln {}_n p_x} {}_n q_x. \end{aligned} \quad (3.62)$$

Z relacji $l_x \cdot {}_n q_x = {}_n d_x$, otrzymujemy

$${}_n K_x = -n \frac{{}_n d_x}{\ln {}_n p_x}. \quad (3.63)$$

Przyrównując otrzymany wynik do prawej strony wzoru (3.15), definiującego czas ekspozycji ${}_n K_x$, tj. do wyrażenia $n \cdot l_{x+n} + {}_n a_x \cdot {}_n d_x$, uzyskujemy równość

$$-n \frac{{}_n d_x}{\ln {}_n p_x} = n \cdot l_{x+n} + {}_n a_x \cdot {}_n d_x. \quad (3.64)$$

Ostatecznie, wzór na współczynnik ${}_n a_x$ ma postać

$$\begin{aligned} {}_n a_x &= -n \frac{l_{x+n}}{{}_n d_x} - \frac{n}{\ln {}_n p_x} = -n \frac{\frac{l_{x+n}}{l_x}}{\frac{{}_n d_x}{l_x}} - \frac{n}{\ln {}_n p_x} = \\ &= -n \frac{{}_n p_x}{{}_n q_x} - \frac{n}{\ln {}_n p_x} = n - \frac{n}{{}_n q_x} - \frac{n}{\ln(1 - {}_n q_x)}. \end{aligned} \quad (3.65)$$

Zbadamy teraz związek pomiędzy współczynnikiem zgonów ${}_n m_x$ oraz prawdopodobieństwem zgonów ${}_n q_x$ w modelu interpolacji wykładniczej. Z definicji cząstkowego współczynnika zgonów mamy

$${}_n m_x = \frac{{}_n d_x}{{}_n K_x}. \quad (3.66)$$

Korzystając z (3.63), wzór (3.66) sprowadza się do postaci

$${}_n m_x = \frac{{}_n Z_x}{{}_n K_x} = \frac{{}_n Z_x}{-n \frac{{}_n Z_x}{\ln {}_n p_x}} = -\frac{1}{n} \ln {}_n p_x = -\frac{1}{n} \ln(1 - {}_n q_x), \quad (3.67)$$

otrzymaliśmy

$${}_n m_x = -\frac{1}{n} \ln(1 - {}_n q_x). \quad (3.68)$$

Relację wiążącą współczynniki i prawdopodobieństwa zgonów można także zapisać równoważnie wzorem

$${}_n q_x = 1 - e^{-n \cdot {}_n m_x}. \quad (3.69)$$

W przypadku szczególnym dla $n = 1$, powyższa formuła redukuje się do postaci

$$q_x = 1 - e^{-m_x}. \quad (3.70)$$

C. Średnie dalsze trwanie życia w modelach interpolacji

Średnie dalsze trwanie życia definiujemy ogólnym wzorem (3.12)

$$e_x = \frac{T_x}{l_x}, \quad (3.71)$$

gdzie l_x jest liczbą dożywających wieku x , a T_x funduszem osobołat życia osób w wieku x ukończonych lat.

Niech T_x^* oznacza fundusz osobołat życia mierzony w pełnych latach. To założenie oznacza, że czas życia osób, które zmarły w danym przedziale wieku, nie jest wliczany do funduszu T_x^* .

Jeśli na początku danego przedziału $[x, x + 1)$ było l_x osób, a do końca tego przedziału dożyło l_{x+1} osób, to łączny czas życia tych osób w podanym przedziale w pełnych latach równy jest liczbie l_{x+1} pomnożonej przez długość przedziału (tutaj 1). Z kolei fundusz osobołat życia T_x^* wyrażony w pełnych latach równy jest sumie pełnych lat przeżytych kolejno w przedziałach $[x, x + 1)$, $[x + 1, x + 2)$ itd. Otrzymujemy więc

$$T_x^* = \sum_{y=x+1}^X l_y = \sum_{k=1}^{X-x} l_{x+k}, \quad (3.72)$$

gdzie X oznacza górną granicę wieku.

Gdyby w obliczeniach uwzględnić dodatkowo czas życia tych osób, które zmarły w przedziałach wieku $[x + k - 1, x + k)$, czyli wyrażenia $a_{x+k-1}d_{x+k-1}$, wówczas fundusz osobołat życia T_x przyjmie postać

$$\begin{aligned} T_x &= \sum_{k=1}^{X-x} (l_{x+k} + a_{x+k-1}d_{x+k-1}) = \\ &= \sum_{k=1}^{X-x} l_{x+k} + \sum_{k=1}^{X-x} a_{x+k-1}(l_{x+k-1} - l_{x+k}). \end{aligned} \quad (3.73)$$

Otrzymujemy w ten sposób formułę na średnie dalsze trwanie życia

$$e_x = \frac{T_x}{l_x} = \sum_{k=1}^{X-x} \frac{l_{x+k}}{l_x} + \sum_{k=1}^{X-x} a_{x+k-1} \frac{l_{x+k-1} - l_{x+k}}{l_x}. \quad (3.74)$$

Rozważymy teraz przypadki szczególne.

Przyjmijmy model interpolacji liniowej, w którym $a_x = \frac{1}{2}$. Wtedy

$$\begin{aligned} T_x &= \sum_{k=1}^{X-x} (l_{x+k} + a_{x+k-1}d_{x+k-1}) = \\ &= \sum_{k=1}^{X-x} l_{x+k} + \frac{1}{2} \sum_{k=1}^{X-x} (l_{x+k-1} - l_{x+k}). \end{aligned} \quad (3.75)$$

Powyższe wyrażenie możemy także przekształcić do postaci

$$T_x = \sum_{k=1}^{X-x} l_{x+k} + \frac{1}{2}(l_x - l_X). \quad (3.76)$$

Zakładamy, że liczba dożywających górnej granicy wieku X wynosi 0, czyli $l_X = 0$. Stąd w modelu interpolacji liniowej e_x wyraża się wzorem

$$\begin{aligned} e_x &= \frac{T_x}{l_x} = \frac{1}{2} + \sum_{k=1}^{X-x} \frac{l_{x+k}}{l_x} = \\ &= \frac{1}{2} + \sum_{k=1}^{X-x} {}_k p_x. \end{aligned} \quad (3.77)$$

Odmienną formułą określamy e_x w modelu interpolacji wykładniczej. Wówczas czas życia tych osób, które zmarły w przedziałach wieku $[x+k-1, x+k)$, czyli wyrażenia $a_{x+k-1}d_{x+k-1}$ są równe

$$a_{x+k-1}d_{x+k-1} = \left(1 - \frac{1}{q_{x+k-1}} - \frac{1}{\ln(1 - q_{x+k-1})}\right) d_{x+k-1}, \quad (3.78)$$

a zatem w modelu interpolacji wykładniczej e_x przyjmuje postać

$$e_x = \sum_{k=1}^{X-x} k p_x + \sum_{k=1}^{X-x} \left(1 - \frac{1}{q_{x+k-1}} - \frac{1}{\ln(1 - q_{x+k-1})}\right) ({}_{k-1}p_x - {}_k p_x). \quad (3.79)$$

Uwagi podsumowujące

1. Formuły dotyczące e_x różnią się w zależności od założonego modelu interpolacyjnego. W każdym jednak przypadku można odwołać się bezpośrednio do definicji e_x jako ilorazu funduszu osobolat T_x i liczby l_x osób dożywających wieku x .
2. Modele interpolacji liniowej i wykładniczej nie są adekwatne dla $x = 0$. Dlatego w przypadku obliczania e_0 wskazane jest uwzględnienie innej metody szacowania współczynnika a_0 .

Przykładowo, w latach 70-tych ubiegłego stulecia Coale i in. zaproponowali, aby przyjąć $a_0 = 0,33$ dla niemowląt płci męskiej oraz $a_0 = 0,35$ dla niemowląt płci żeńskiej w populacji o wysokim współczynniku zgonów niemowląt, tj. gdy $m_0 \geq 0,107$. W innych przypadkach wartość a_0 uzależnili od wartości m_0 za pomocą relacji $a_0 = 0,045 + 2,684 \cdot m_0$ dla niemowląt płci męskiej oraz $a_0 = 0,053 + 2,800 \cdot m_0$ dla niemowląt płci żeńskiej.

3. W metodologii GUS odmienny jest sposób obliczania osobolat życia K_0 . W przypadku grupy wieku $[0, 1)$ lat, zamiast ogólnej formuły (3.34), korzysta się z relacji (3.38).

3.6. TTŻ w analizie czasu trwania zjawisk

Tablice trwania życia mogą być wykorzystywane nie tylko do analizy umieralności *sensu stricto* lub, innymi słowy, do opisu czasu trwania życia jednostek należących do danej populacji, ale mogą też znaleźć zastosowanie w analizie czasu trwania dowolnych zjawisk. Niektóre przykłady zastosowań TTŻ zawiera tablica 11.

Tablica 11. Przykłady zastosowań TTŻ

Nazwa zjawiska	Jednostki badane	Zdarzenie inicjujące	Zdarzenie kończące	Czas własny
umieralność	osoby żyjące	urodzenie	zgon	wiek
zawieranie pierwszego małżeństwa	osoby w stanie wolnym	urodzenie	zawarcie pierwszego małżeństwa	wiek
urodzenie pierwszego dziecka w małżeństwie	małżeństwa bezdzietne	zawarcie małżeństwa	urodzenie pierwszego dziecka	staż związku małż. bez dzieci
staż małżeński	pary w związku małżeńskim	zawarcie małżeństwa	rozpad małżeństwa zgon/rozwód	czas trwania małżeństwa
migracja z miejsca urodzenia	osoby zam. w miejscu urodzenia	urodzenie	zmiana miejsca zamieszkania	czas pobytu (wiek)
rozpoczęcie pracy zawodowej	osoby, które nigdy nie pracowały	urodzenie (ukończenie edukacji)	rozpoczęcie pierwszej pracy	wiek (czas do podjęcia pracy)
trwanie bezrobocia	osoby bezrobotne	utrata pracy	rozpoczęcie pracy	czas poszukiwania pracy

Źródło: opracowanie własne

3.7. Koncepcja populacji stacjonarnej

Rozważymy populację zamkniętą (bez migracji), w której ruch rzeczywisty określają jedynie urodzenia i zgony. Przyjmijmy dodatkowo:

- stałą liczbę urodzeń $U_0/365$ dziennie, tj. U_0 rocznie,
- stałe prawdopodobieństwa ${}_x p_0$ – jednakowe dla dziennych kohort urodzeniowych.

Populacja spełniająca wymienione warunki nosi miano **populacji stacjonarnej** lub zamiennie – **populacji zastojowej**.

Przykład 14. Załóżmy, że dla dziennych kohort urodzeniowych odnotowanych w ciągu roku w pewnej zamkniętej populacji spełnione są założenia populacji zastojowej. Proces wymierania tych kohort w kolejnych dniach roku ilustruje tablica 12.

Tablica 12. Proces wymierania dziennych kohort urodzeniowych
w ciągu roku

wiek $x(\text{dni})$	${}_x p_0$	Populacja żyjących w dniu:							
		1.01	2.01	3.01	4.01	5.01	6.01	...	31.12
0	1,000	1000	1000	1000	1000	1000	1000	...	1000
1	0,970		970	970	970	970	970	...	970
2	0,950			950	950	950	950	...	950
3	0,946				946	946	946	...	946
4	0,942					942	942	...	942
...
364	0,900								900

Źródło: dane umowne

Jeśli własności te utrzymują się niezmiennie od wielu lat, wówczas tablicę 12 możemy uzupełnić do postaci prezentowanej w tablicy 13.

Tablica 13. Ilustracja populacji stacjonarnej

wiek $x(\text{dni})$	${}_x p_0$	Populacja żyjących w dniu:							
		1.01	2.01	3.01	4.01	5.01	6.01	...	31.12
0	1,000	1000	1000	1000	1000	1000	1000	...	1000
1	0,970	970	970	970	970	970	970	...	970
2	0,950	950	950	950	950	950	950	...	950
3	0,946	946	946	946	946	946	946	...	946
4	0,942	942	942	942	942	942	942	...	942
...
364	0,900	900	900	900	900	900	900	...	900

Źródło: dane umowne

Z analizy danych w tablicy 13 płynie kilka spostrzeżeń, które uogólnić można na przypadek dowolnej populacji stacjonarnej:

1. Zauważmy, że liczba osób, które w dniu $\tau = 31.12$ nie ukończyły 1-go roku życia jest równa sumie wyrazów w ostatniej kolumnie tablicy 13, czyli

$$L_0(\tau) = L_0 = 1000 + 970 + 950 + 946 + 942 + \dots + 900.$$

Suma ta jest jednakowa dla dowolnego τ .

Uogólniając, w populacji stacjonarnej liczba osób $L_x(\tau)$ w wieku x w chwili τ jest stała w czasie, tj. $L_x(\tau) = L_x$ dla dowolnego τ .

2. Oznacza to, że stała w czasie jest również łączna liczebność populacji $L(\tau)$, bowiem stałe są elementy składowe $L_0(\tau), L_1(\tau), \dots$. Mamy więc

$$L(\tau) = L = L_0 + L_1 + L_2 + L_3 + \dots,$$

a tym samym stała jest też struktura populacji wg wieku, wyrażona wskaźnikami struktury

$$\frac{L_0}{L}, \frac{L_1}{L}, \frac{L_2}{L}, \dots$$

3. Liczba osobołat życia w ciągu roku dla pojedynczej, dziennej kohorty urodzeniowej z dowolnego dnia, np. z dnia 1 stycznia, obliczona na podstawie danych z tablicy 13 wynosi

$$\frac{1}{365} \cdot 1000 + \frac{1}{365} \cdot 970 + \frac{1}{365} \cdot 950 + \frac{1}{365} \cdot 946 + \frac{1}{365} \cdot 942 + \dots + \frac{1}{365} \cdot 900.$$

Mnożąc przez łączną liczbę dziennych kohort, otrzymamy liczbę osobołat życia K_0 rocznej kohorty urodzeniowej (licząc od dnia urodzenia przez okres 1 roku), czyli

$$K_0 = 365 \cdot \left[\frac{1}{365} (1000 + 970 + 950 + 946 + 942 + \dots + 900) \right].$$

Z porównania wyniku z liczebnością L_0 otrzymujemy równość

$$K_0 = L_0.$$

Podobne równości zachodzą także dla dowolnego x

$$K_1 = L_1, \quad K_2 = L_2, \quad K_3 = L_3, \dots$$

Uogólniając, liczba osobołat życia K_x w ciągu roku rocznej kohorty urodzeniowej dożywającej wieku x lat równa jest liczebności $L_x(\tau)$.

4. Z powyższego wynika, że łączna liczba osobołat życia K dla rocznej kohorty urodzeniowej wynosi

$$K = L_0 + L_1 + L_2 + \dots$$

5. Stały w czasie jest też porządek wymierania kolejnych generacji, co oznacza, że liczba zgonów I rodzaju jest równa liczbie zgonów III rodzaju.

6. Fundusz T_0 osobolet życia (będący sumą $K_0 + K_1 + K_2 + \dots$) jest równy ogólnej liczbie żyjących w populacji w tym samym momencie

$$T_0 = L_0 + L_1 + L_2 + \dots$$

Analogiczne relacje zachodzą dla dowolnego x .

7. Skoro liczebność populacji jest stała w czasie, to liczba zgonów D w danym okresie t równa się liczbie urodzeń U_0 w tym okresie. W konsekwencji, ogólny przekrojowy współczynnik zgonów CDR oraz ogólny przekrojowy współczynnik urodzeń CBR są jednakowe, czyli

$$CDR = CBR.$$

8. Przekrojowe współczynniki urodzeń i zgonów równe są odwrotności średniego trwania życia noworodka e_0 . Z definicji bowiem e_0 jest ilorazem funduszu osobolet T_0 życia kohorty urodzeniowej do jej liczebności początkowej U_0 . Liczebność tę możemy obliczyć, sumując liczby zgonów D_x w kolejnych grupach wieku (w ujęciu kohortowym) lub jako sumę liczby zgonów w zadanym okresie w zbiorowości ogółu żyjących (ujęcie przekrojowe). Mamy więc

$$U_0 = \sum_{x=0}^{X-1} D_x = D,$$

a ponadto

$$T_0 = K.$$

Otrzymujemy

$$e_0 = \frac{T_0}{U_0} = \frac{K}{D} = \frac{K}{B},$$

czyli

$$e_0 = \frac{1}{CDR} = \frac{1}{CBR}$$

lub równoważnie

$$CDR = CBR = \frac{1}{e_0}.$$

9. W przypadku modelu ludności stacjonarnej oczekiwany czas życia noworodka e_0 oraz średni wiek \bar{x} jednostek w chwili zgonu są równe. Równość $e_0 = \bar{x}$ w ogólnym przypadku nie zachodzi [zob. sekcja 6.5.6 w rozdziale 6].

Uwagi podsumowujące

1. Model populacji stacjonarnej (zastojowej) pozwala powiązać ze sobą główne parametry tablicy trwania życia, takie jak średnie dalsze trwanie życia, współczynniki urodzeń, zgonów czy czas ekspozycji. Umożliwia to estymowanie wybranych parametrów tablicowych na podstawie innych parametrów.
2. Model ten znajduje zastosowanie np. w demografii historycznej. Przyjmuje się, że populacje żyjące w dawnych wiekach posiadały własności zbliżone do populacji stacjonarnych (ze względu na ich powolny wzrost liczebny). Przykładowo, analiza liczby prehistorycznych szkieletów z określonej epoki i przyporządkowanie ich do odpowiednich grup wieku może być podstawą do szacowania współczynnika urodzeń.
3. Model ludności stacjonarnej może być użyty do wyznaczania parametrów populacji zamkniętej, dla której zdefiniowano pojęcie „urodzenia”, „zgonu”, a także „czasu życia”. Przykładem jest zbiorowość uczniów, w której „urodzeniem” i „zgonem” są odpowiednio rozpoczęcie i ukończenie szkoły, a „czasem życia” – czas trwania nauki.

Ćwiczenia

Ćwiczenie 3.1

Dany jest fragment skróconej tablicy trwania życia (tablica 14). Obliczyć:

1. Liczbę zgonów w grupie wieku [20, 30) lat oraz [50, 60) lat,
2. Współczynnik zgonów w grupie wieku [20, 30) lat oraz [50, 60) lat,
3. Prawdopodobieństwo zgonu przed upływem 10 lat dla 20-latków,
4. Prawdopodobieństwo przeżycia 10 lat w grupie 20-latków,
5. Prawdopodobieństwo zgonu przed upływem 10 lat dla 50-latków,
6. Prawdopodobieństwo przeżycia 10 lat w grupie 50-latków,
7. Prawdopodobieństwo zgonu przed upływem 20 lat dla 70-latków,
8. Prawdopodobieństwo przeżycia 20 lat w grupie 70-latków,
9. Średnie dalsze trwanie życia noworodka,

10. Średnie dalsze trwanie życia osób w wieku 30 ukończonych lat,
 11. Średnie dalsze trwanie życia osób w wieku 50 ukończonych lat.

Tablica 14. Fragment skróconej tablicy trwania życia

Wiek x (w latach)	l_x (w tys.)	${}_{10}K_x$ (w tys.)	T_x (w tys.)
0	100,0	997,5	8117
10	99,5	993,5	7119,5
20	99,2	991	6126
30	99,0	985	5135
40	98,0	975	4150
50	97,0	945	3175
60	92,0	870	2230
70	82,0	720	1360
80	62,0	430	640
90	24,0	165	210
100	9,0	45	45
110	0,0	0	0

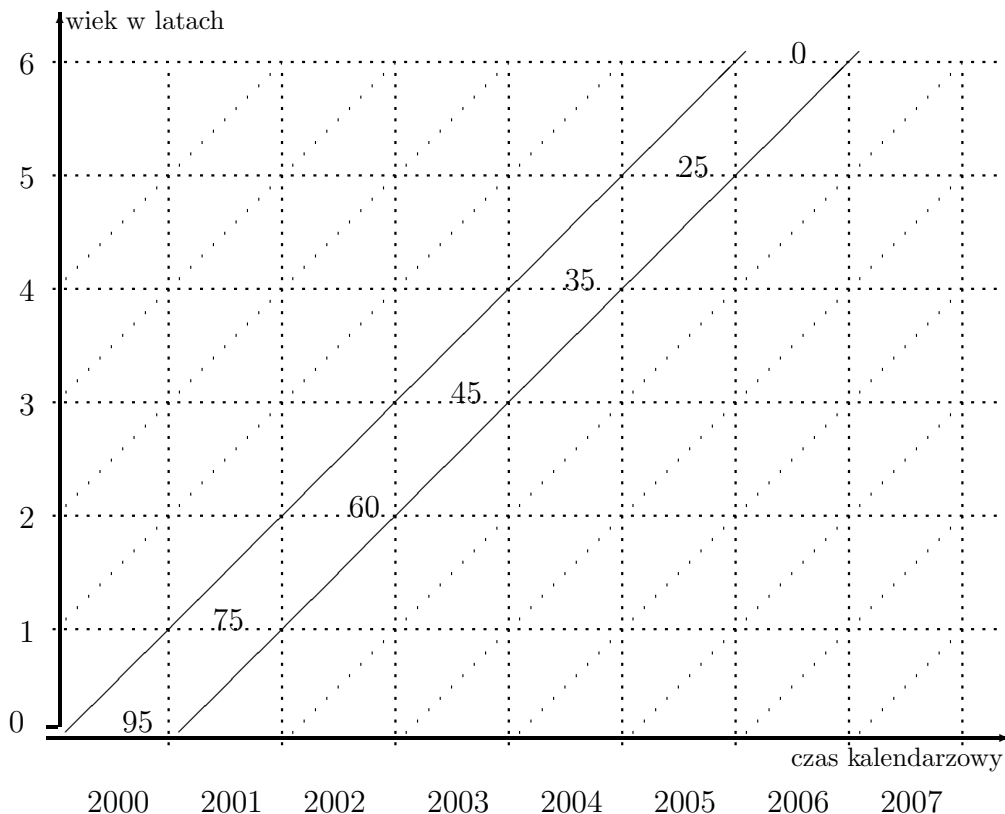
Źródło: dane umowne

Ćwiczenie 3.2

Dany jest fragment siatki demograficznej (rysunek 27), przedstawiający liczbę osób urodzonych w roku 2000 i liczby dożywających I rodzaju (w tys.).

Korzystając z danych zamieszczonych w siatce, zbudować tablicę trwania życia (tj. wyznaczyć $l_x, d_x, q_x, p_x, T_x, e_x$), wykorzystując model interpolacji:

1. Liniowej,
2. Wykładniczej.



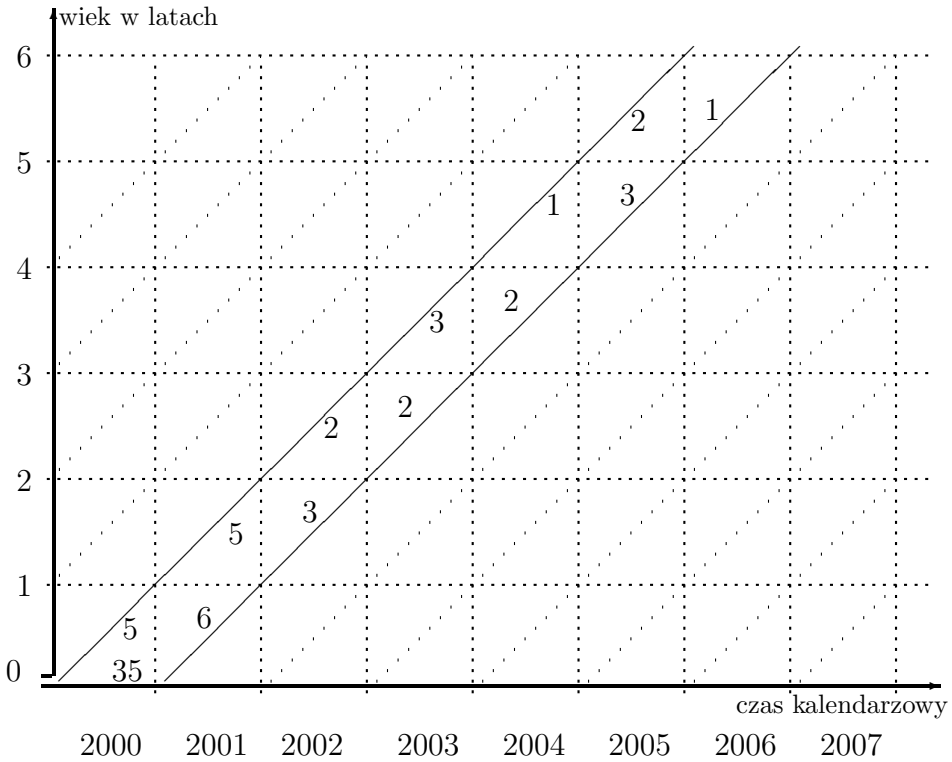
Rysunek 27. Fragment siatki demograficznej oraz liczby dożywających I rodzaju (dane umowne)

Ćwiczenie 3.3

Dany jest fragment siatki demograficznej (rysunek 28), przedstawiający liczebność początkową pewnej generacji oraz liczby zgonów młodszych i starszych (w tys.).

Na podstawie danych zamieszczonych w siatce zbudować tablicę trwania życia (tj. wyznaczyć $l_x, d_x, q_x, p_x, T_x, e_x$), stosując model interpolacji:

1. Liniowej,
2. Wykładniczej.



Rysunek 28. Fragment siatki demograficznej oraz liczby zgonów młodszych i starszych (dane umowne)

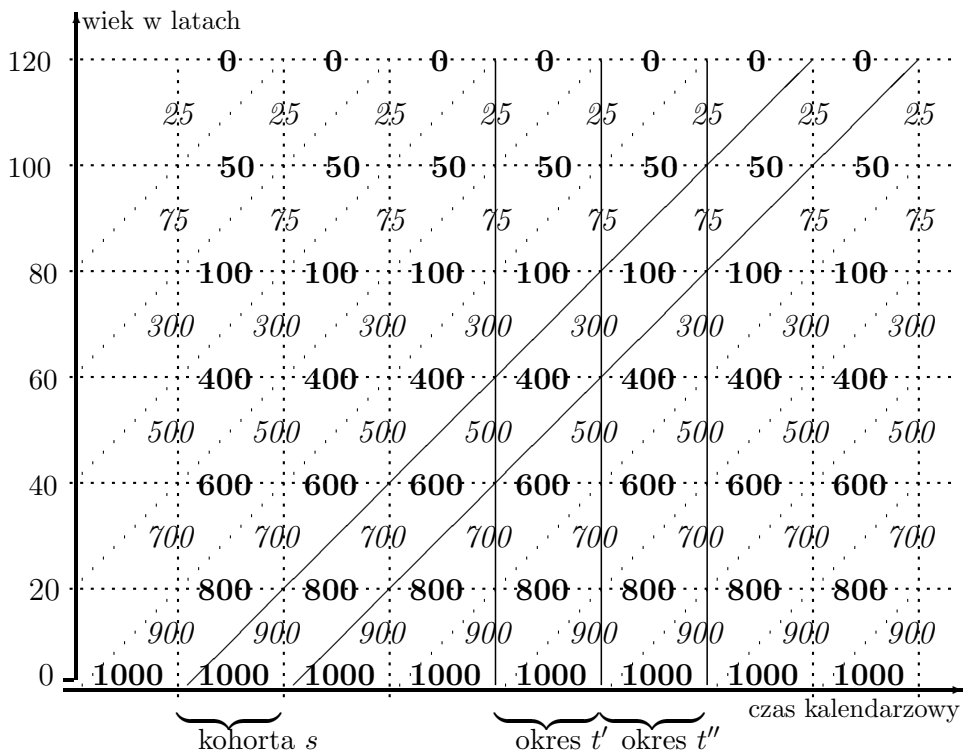
Ćwiczenie 3.4

Rozważmy przykładową populację zastożową, w której porządek wymierania przedstawia rysunek 29. Liczby dożywających I rodzaju (w tys.) zaznaczono na wykresie drukiem pogrubionym, natomiast liczby dożywających II rodzaju (w tys.) – drukiem pochylonym. Przyjętą jednostką czasu jest 20 lat.

Ustalić i porównać:

1. Liczby zgonów I rodzaju w poszczególnych 20-letnich grupach wieku w kohorcie s oraz liczby zgonów III rodzaju w analogicznych grupach wieku w okresach t' oraz t'' ,
2. Liczby dożywających I rodzaju osiągających wiek 40 lat w dwóch dowolnych kohortach,
3. Liczebność populacji we wszystkich grupach wieku łącznie na koniec okresu t' oraz t'' ,

4. Strukturę ludności wg grup wieku na początku i końcu okresu t' oraz t'' ,
5. Łączną liczbą zgonów i liczbę urodzeń w dowolnym okresie kalendarzowym.



Rysunek 29. Siatka demograficzna i liczby dożywających I i II rodzaju (dane umowne)

Rozdział 4

ROZKŁADY CZASU TRWANIA ŻYCIA

4.1. Wprowadzenie

W rozważaniach dotyczących tablic trwania życia poznaliśmy m.in. ogólny model tablicy trwania życia oraz jego zastosowanie do konstrukcji tablic kohortowych i przekrojowych.

Ujęcie tablicowe ma tę wadę, iż proces wymierania danej kohorty urodzeniowej jest przedstawiony w TTŻ w podziale na arbitralne przedziały wieku $[x, x+n)$, gdzie x, n to liczby nieujemne, całkowite. Jednak w niektórych zastosowaniach, np. w ubezpieczeniach na życie, ważna staje się możliwość wyznaczenia wybranych charakterystyk czasu trwania życia dla dowolnego $x \geq 0$ lub dla dowolnego przedziału $[x, x+y)$, $y > 0$. Ujęcie takie jest możliwe dzięki traktowaniu czasu życia jako zmiennej losowej o pewnym ciągłym rozkładzie prawdopodobieństwa. To podejście zostanie przedstawione w dalszej części tego rozdziału.

4.2. Niektóre charakterystyki funkcyjne i liczbowe

Definicja 1. Niech X będzie nieujemną i ciągłą zmienną losową reprezentującą czas życia noworodka. Niech $F_X(x)$ będzie dystrybuantą rozkładu zmiennej X , czyli

$$F_X(x) = P(X < x), \quad (4.1)$$

taką, że $F_X(0) = 0$.

Funkcją przeżycia S_X zmiennej X nazywamy funkcję komplementarną do F_X , tj.

$$S_X(x) = 1 - F_X(x). \quad (4.2)$$

Będziemy pisać dalej $F(x)$, $S(x)$ zamiast $F_X(x)$, $S_X(x)$ w przypadkach, gdy nie będzie wątpliwości, o jaką zmienną losową chodzi.

Z definicji 1 wynika, że dla ustalonego $x \geq 0$ wartość $S_X(x)$ w punkcie x określa prawdopodobieństwo zdarzenia, że noworodek dożyje wieku x .

Definicja 2. Załóżmy, że $x \geq 0$ jest ustaloną liczbą rzeczywistą. Niech

$$T(x) = X - x \quad \text{dla} \quad X \geq x. \quad (4.3)$$

Zmienną $T(x)$ nazywamy pozostałym czasem życia osoby w wieku x .

Dystrybuanta zmiennej losowej $T(x)$ dla zadanego $y \geq 0$ wyraża się wzorem

$$\begin{aligned} F_{T(x)}(y) &= P(T(x) < y) = \\ &= P(X - x < y \mid X \geq x) = P(X < x + y \mid X \geq x) = \\ &= 1 - P(X \geq x + y \mid X \geq x) = \\ &= 1 - \frac{P(X \geq x + y)}{P(X \geq x)} = 1 - \frac{S_X(x + y)}{S_X(x)}. \end{aligned} \quad (4.4)$$

Zauważymy, że dla $x = 0$ mamy $T(0) = X$ i $F_{T(0)} = F_X$, czyli zmienna X jest szczególnym przypadkiem zmiennej losowej $T(x)$.

Niech ${}_y p_x$ oznacza prawdopodobieństwo zdarzenia, że osoba w wieku x lat przeżyje co najmniej y następnych lat, natomiast ${}_y q_x$ prawdopodobieństwo, że osoba w wieku x lat umrze przed upływem y lat. Prawdopodobieństwa ${}_y p_x$ oraz ${}_y q_x$ są wyrażone wzorami

$${}_y p_x = \frac{S_X(x + y)}{S_X(x)}, \quad (4.5)$$

$${}_y q_x = 1 - \frac{S_X(x + y)}{S_X(x)}, \quad (4.6)$$

przy czym zachodzi oczywisty związek ${}_y p_x + {}_y q_x = 1$. Gdy $x = 0$ mamy także ${}_y p_0 = S_X(y)$ oraz ${}_y q_0 = 1 - S_X(y) = F_X(y)$.

W przypadku gdy $y = 1$, zamiast ${}_y p_x, {}_y q_x$ będziemy pisać p_x, q_x .

Pokazaliśmy w (4.4), że dystrybuanta zmiennej $T(x)$ w punkcie $y \geq 0$ ma postać

$$F_{T(x)}(y) = 1 - \frac{S_X(x+y)}{S_X(x)}, \quad (4.7)$$

czyli funkcja $S_{T(x)}$ komplementarna do $F_{T(x)}$ wyraża się wzorem

$$S_{T(x)}(y) = 1 - F_{T(x)}(y) = \frac{S_X(x+y)}{S_X(x)}. \quad (4.8)$$

Z porównania prawej strony (4.7) z (4.6) oraz prawej strony (4.8) z (4.5) wynika, że dystrybuanta rozkładu reszty życia osoby w wieku x równa jest prawdopodobieństwu ${}_yq_x$, a funkcja komplementarna do tej dystrybuanty – prawdopodobieństwu ${}_yp_x$, czyli dla ustalonego x

$$F_{T(x)}(y) \equiv {}_yq_x, \quad S_{T(x)}(y) \equiv {}_yp_x. \quad (4.9)$$

Definicja 3. Niech $F'_X(x)$ oznacza pochodną dystrybuanty $F_X(x)$, tj. wartość funkcji gęstości f_X zmiennej X w punkcie x . Intensywnością (natężeniem) zgonów nazywamy funkcję

$$\mu_x = \frac{f_X(x)}{S_X(x)}, \quad \text{gdym } S_X(x) > 0. \quad (4.10)$$

Iloczyn $\mu_x dx$ oznacza w przybliżeniu prawdopodobieństwo zgonu w bardzo małym przedziale wieku $[x, x + dx)$ w grupie osób, które dożyły wieku x .

Zauważymy, że na mocy (4.10) oraz (4.5) funkcję gęstości $f_X(x)$ można przedstawić w postaci iloczynu

$$f_X(x) = {}_xp_0 \cdot \mu_x, \quad x \geq 0. \quad (4.11)$$

Po scałkowaniu stronami (4.10) na przedziale $[0, x]$, otrzymamy

$$\int_0^x \mu_v dv = \int_0^x \frac{f_X(v)}{S_X(v)} dv. \quad (4.12)$$

Dokonyamy zamiany zmiennych pod całką. Niech $z = S_X(v)$. Mamy wtedy $dS_X(v) = dz$ lub inaczej $S'_X(v)dv = dz$. Ponieważ $S_X(v) =$

$1 - F_X(v)$, więc pochodna funkcji $S_X(v)$ względem v równa jest $-f_X(v)$, wobec tego mamy także $-f_X(v)dv = dz$. Otrzymujemy

$$\int_0^x \mu_v dv = - \int_1^{S_X(x)} \frac{1}{z} dz = - \ln z \Big|_1^{S_X(x)} = - \ln S_X(x). \quad (4.13)$$

Z powyższego wyniku więc następujący związek

$$S_X(x) = e^{-\int_0^x \mu_v dv}, \quad (4.14)$$

tym samym

$$F_X(x) = 1 - e^{-\int_0^x \mu_v dv}. \quad (4.15)$$

W innym zapisie mamy

$$S_X(x) = \exp \left\{ - \int_0^x \mu_v dv \right\} \quad (4.16)$$

oraz

$$F_X(x) = 1 - \exp \left\{ - \int_0^x \mu_v dv \right\}, \quad (4.17)$$

stąd dostajemy również

$$F_{T(x)}(y) = 1 - \frac{S_X(x+y)}{S_X(x)} = 1 - \frac{\exp \left\{ - \int_0^{x+y} \mu_v dv \right\}}{\exp \left\{ - \int_0^x \mu_v dv \right\}} = \quad (4.18)$$

$$= 1 - \exp \left\{ - \int_x^{x+y} \mu_v dv \right\}$$

oraz

$$S_{T(x)}(y) = 1 - F_{T(x)}(y) = \exp \left\{ - \int_x^{x+y} \mu_v dv \right\}. \quad (4.19)$$

Korzystając z (4.4) oraz (4.10), obliczymy na koniec gęstość $f_{T(x)}$ zmiennej losowej $T(x)$. Ponieważ $F_{T(x)}(y)$ wyraża się wzorem (4.7), więc pochodna funkcji $F_{T(x)}(y)$ względem $y \geq 0$ jest równa

$$\begin{aligned} f_{T(x)}(y) &= \frac{dF_{T(x)}(y)}{dy} = - \frac{1}{S_X(x)} \frac{dS_X(x+y)}{dy} = \\ &= - \frac{1}{S_X(x)} (-f_X(x+y)) = \frac{f_X(x+y)}{S_X(x)} = \quad (4.20) \\ &= \frac{S_X(x+y)}{S_X(x)} \cdot \frac{f_X(x+y)}{S_X(x+y)} = {}_y p_x \cdot \mu_{x+y}. \end{aligned}$$

W przypadku szczególnym, gdy $x = 0$, funkcja gęstości zmiennej $T(x)$ sprowadza się do funkcji gęstości (4.11) zmiennej X , czyli do

$$f_X(x) = {}_x p_0 \cdot \mu_x. \quad (4.21)$$

Zauważymy, że funkcja intensywności μ_x jednoznacznie determinuje zarówno dystrybuantę, jak i gęstość rozkładu czasu życia noworodka, tj. zmiennej losowej X , a także czasu życia osoby w wieku x , reprezentowanego przez zmienną losową $T(x)$. Funkcje te można wyznaczyć na podstawie intensywności zgonów μ_x , korzystając ze związków (4.14)–(4.19) oraz (4.20)–(4.21).

Definicja 4. *Wartość oczekiwana czasu trwania życia, tj. zmiennej losowej $X \geq 0$ o rozkładzie ciągłym z gęstością $f_X(x)$, jest określona wzorem*

$$E(X) = \int_0^{\infty} x \cdot f_X(x) dx. \quad (4.22)$$

$E(X)$ nazywać będziemy średnim trwaniem życia noworodka i oznaczać symbolem \dot{e}_0 .

Ponieważ funkcję gęstości $f_X(x)$ możemy wyrazić za pomocą iloczynu $S(x) \cdot \mu_x$, wobec tego wartość oczekiwaną czasu życia noworodka możemy zapisać także w postaci

$$\dot{e}_0 \equiv E(X) = \int_0^{\infty} x \cdot S(x) \cdot \mu_x dx. \quad (4.23)$$

Po dokonaniu całkowania przez części prawej strony (4.23) oraz po uwzględnieniu faktu, że $x \cdot S(x) \rightarrow 0$, gdy $x \rightarrow \infty$ oraz $x \cdot S(x) = 0$, gdy $x = 0$, mamy

$$\dot{e}_0 \equiv E(X) = -x \cdot S(x) \Big|_0^{\infty} + \int_0^{\infty} S(x) dx = \int_0^{\infty} S(x) dx. \quad (4.24)$$

Pokażemy, że wartość oczekiwana (4.24) jest tożsama w tzw. modelu ciągłym ze średnim trwaniem życia noworodka (3.12) z tablic trwania życia, zdefiniowanym jako iloraz T_0/l_0 , przy założeniu, że liczby dożywających do określonego wieku są funkcjami ciągłymi względem wieku.

Niech l_{x+y} dla $y \in [0, 1]$ oznacza liczbę dożywających wieku $x + y$ lat. Założymy, że l_{x+y} jest ciągłą funkcją zmiennej $y \in [0, 1]$. Wówczas

czas ekspozycji K_x na ryzyko zgonu w grupie wieku $[x, x + 1)$ możemy zapisać jako całkę

$$K_x = \int_0^1 l_{x+y} dy, \quad (4.25)$$

natomiast fundusz T_0 osobołat życia, zdefiniowany przez analogię do (3.7), zapiszemy za pomocą następującej sumy całek

$$T_0 = \sum_{u=0}^{\infty} K_u = \int_0^1 l_{0+y} dy + \int_0^1 l_{1+y} dy + \int_0^1 l_{2+y} dy + \dots \quad (4.26)$$

Po dokonaniu odpowiedniej zamiany zmiennych pod poszczególnymi całkami powyższa suma daje się sprowadzić do postaci

$$T_0 = \int_0^1 l_x dx + \int_1^2 l_x dx + \int_2^3 l_x dx + \dots = \int_0^{\infty} l_x dx. \quad (4.27)$$

Ostatecznie, na podstawie (3.12), (3.11) oraz (4.27) mamy

$$\frac{T_0}{l_0} = \frac{1}{l_0} \int_0^{\infty} l_x dx = \int_0^{\infty} {}_x p_0 dx = \int_0^{\infty} S(x) dx. \quad (4.28)$$

Otrzymaliśmy wartość oczekiwaną zmiennej losowej X , tj. (4.24).

Można pokazać w drodze podobnego rozumowania i przy analogicznym założeniu, że

$$\frac{T_x}{l_x} = E(T(x)) \equiv \dot{e}_x, \quad \text{dla } x = 1, 2, \dots \quad (4.29)$$

4.3. Intensywność zgonów a współczynnik zgonów

Cząstkowy, tablicowy współczynnik zgonów (3.16), przedstawiony w rozdziale 3, ma postać

$${}_n m_x = \frac{{}_n d_x}{{}_n K_x}, \quad (4.30)$$

gdzie ${}_n d_x$ oraz ${}_n K_x$ oznaczają odpowiednio liczbę zgonów i czas ekspozycji na ryzyko zgonu w przedziale wieku $[x, x + n)$ lat. Warto dodatkowo przypomnieć, że liczba zgonów ${}_n d_x$ jest równa różnicy liczb dożywających wieku x oraz $x + n$, czyli

$${}_n d_x = l_x - l_{x+n}, \quad (4.31)$$

z kolei czas ekspozycji ${}_nK_x$ na ryzyko zgonu w przedziale $[x, x+n)$ wyraża się wzorem

$${}_nK_x = n \cdot l_{x+n} + {}_nq_x \cdot {}_nd_x. \quad (4.32)$$

Zbadamy granicę $\lim_{n \rightarrow 0} {}_nm_x$. Zauważymy, że dla małych n czas ekspozycji ${}_nK_x$ może być przybliżony jako

$${}_nK_x \approx n \cdot l_x. \quad (4.33)$$

Przybliżenie to jest tym lepsze, im mniejsza długość n przedziału.

Dla $n \rightarrow 0$ mamy

$$\lim_{n \rightarrow 0} {}_nm_x = \lim_{n \rightarrow 0} \frac{l_x - l_{x+n}}{n \cdot l_x}. \quad (4.34)$$

Korzystając z definicji pochodnej funkcji, stwierdzamy, że

$$\lim_{n \rightarrow 0} \frac{l_x - l_{x+n}}{n} = -l'_x, \quad (4.35)$$

a stąd otrzymujemy

$$\lim_{n \rightarrow 0} {}_nm_x = -\frac{l'_x}{l_x}. \quad (4.36)$$

Ponieważ $l_x = l_0 \cdot S(x)$, więc mamy dalej

$$\lim_{n \rightarrow 0} {}_nm_x = -\frac{l_0 \cdot S'(x)}{l_0 \cdot S(x)} = \frac{f_X(x)}{S(x)} = \mu_x, \quad (4.37)$$

co oznacza, że intensywność zgonów μ_x jest granicą, do której dążą współczynniki zgonów ${}_nm_x$, gdy $n \rightarrow 0$.

4.4. Relacja między ${}_yq_x$ a prawdopodobieństwami tablicowymi q_x

Związek pomiędzy prawdopodobieństwami ${}_yq_x$ oraz q_x omówimy dla przypadku, w którym x, y są liczbami nieujemnymi, całkowitymi [Skałba 2002].

Niech $x_1 < x_2 < \dots < x_X$ będzie ustalonym podziałem osi czasu własnego $[0, X]$ (wieku). Przyjmijmy dla uproszczenia, że $x_j = j$, tj.

$$x_0 = 0, \quad x_1 = 1, \quad x_2 = 2, \quad x_3 = 3, \quad \dots, \quad x_X = X. \quad (4.38)$$

Wtedy

$$\begin{aligned}
 {}_j p_i &= \frac{S(i+j)}{S(i)} = \frac{S(i+1)}{S(i)} \cdot \frac{S(i+2)}{S(i+1)} \cdot \dots \cdot \frac{S(i+j)}{S(i+j-1)} = \\
 &= p_i \cdot p_{i+1} \cdot \dots \cdot p_{i+j-1} = \prod_{k=1}^j p_{i+k-1} = \quad (4.39)
 \end{aligned}$$

$$= (1 - q_i) \cdot (1 - q_{i+1}) \cdot \dots \cdot (1 - q_{i+j-1}) = \prod_{k=1}^j (1 - q_{i+k-1}).$$

Otrzymaliśmy równość

$${}_j p_i = \prod_{k=1}^j (1 - q_{i+k-1}), \quad (4.40)$$

co prowadzi także do równości

$${}_j q_i = 1 - {}_j p_i = 1 - \prod_{k=1}^j (1 - q_{i+k-1}). \quad (4.41)$$

Z drugiej strony mamy także na podstawie wzorów (4.5) oraz (4.6)

$${}_j p_i = \frac{S(i+j)}{S(i)} = \frac{l_{i+j}}{l_i}, \quad (4.42)$$

$${}_j q_i = 1 - \frac{S(i+j)}{S(i)} = 1 - \frac{l_{i+j}}{l_i}, \quad (4.43)$$

gdzie l_j jest parametrem tablicy trwania życia oznaczającym liczbę dożywających wieku $x_j = j$.

Przykład 15. Obliczymy prawdopodobieństwo zgonu w ciągu 3 lat w grupie osób (mężczyzn bądź kobiet), które dożyły wieku 25 lat, korzystając z danych zamieszczonych w tablicach 9 i 10.

I sposób: na podstawie prawdopodobieństw tablicowych q_x

Skorzystamy najpierw z relacji (4.41) oraz z danych dotyczących prawdopodobieństw q_x zawartych w tablicy trwania życia dla mężczyzn (tablica 9). Mamy

$$\begin{aligned} {}_3q_{25}^{(M)} &= 1 - \prod_{k=1}^3 \left(1 - q_{25+k-1}^{(M)} \right) = \\ &= 1 - (1 - q_{25}^{(M)}) \cdot (1 - q_{26}^{(M)}) \cdot (1 - q_{27}^{(M)}) = \\ &= 1 - (1 - 0,00097) \cdot (1 - 0,00101) \cdot (1 - 0,00105) = \\ &= 1 - 0,996973 \approx 0,00303. \end{aligned}$$

Podobnie, w przypadku kobiet skorzystamy z relacji (4.41) oraz z danych zawartych w tablicy 10, czyli

$$\begin{aligned} {}_3q_{25}^{(K)} &= 1 - \prod_{k=1}^3 \left(1 - q_{25+k-1}^{(K)} \right) = \\ &= 1 - (1 - q_{25}^{(K)}) \cdot (1 - q_{26}^{(K)}) \cdot (1 - q_{27}^{(K)}) = \\ &= 1 - (1 - 0,00023) \cdot (1 - 0,00025) \cdot (1 - 0,00026) = \\ &= 1 - 0,99926 \approx 0,00074. \end{aligned}$$

II sposób: na podstawie liczb dożywających l_x

Tym razem skorzystamy z relacji (4.43) oraz z liczb dożywających do określonego wieku dla mężczyzn i kobiet, zawartych w tablicach odpowiednio 9 i 10. Mamy

$${}_3q_{25}^{(M)} = 1 - \frac{l_{28}^{(M)}}{l_{25}^{(M)}} = 1 - \frac{98417}{98716} \approx 0,00303,$$

$${}_3q_{25}^{(K)} = 1 - \frac{l_{28}^{(K)}}{l_{25}^{(K)}} = 1 - \frac{99196}{99269} \approx 0,00074.$$

Podsumowując wyniki uzyskane w przykładzie 15, warto zauważyć, że prawdopodobieństwo zgonu w ciągu 3 lat dla 25-latków jest ponad czterokrotnie wyższe w przypadku mężczyzn niż kobiet.

4.5. Prawa umieralności – ujęcie historyczne

W literaturze znanych jest wiele prób modelowania intensywności zgonów, prawdopodobieństw zgonów lub prawdopodobieństw przeżycia w określonych populacjach. Najprostszym ujęciem jest model wykładniczy, w którym intensywność zgonów jest funkcją stałą

$$\mu_x \equiv \mu = \text{const}, \quad x \geq 0. \quad (4.44)$$

Wzór (4.44) rzadko jednak znajduje zastosowanie w modelowaniu umieralności generacji ludzkich, ze względu na nierealistyczne założenie o jednakowym natężeniu zgonów dla dowolnego wieku x .

W modelach demograficznych zakłada się niekiedy, że μ_x jest funkcją przedziałami stałą (*piecewise exponential model*), co dla dostatecznie wąskich przedziałów wieku może być dobrym przybliżeniem rzeczywistej intensywności zgonów. Przedziałami wykładniczy model umieralności sprowadza się wówczas do modelu interpolacji, przedstawionego w sekcji 3.5.2.

Do historycznych praw umieralności zaliczamy model de Moivre'a, Lamberta, Gompertza-Makehama czy model Weibulla [Frątczak 1997].

Najstarszy model de Moivre'a z roku 1725 przedstawia natężenie wymierania jako funkcję wieku x postaci

$$\mu_x = \frac{1}{X - x} \quad \text{dla} \quad 0 < x < X, \quad (4.45)$$

gdzie X oznacza nieprzekraczalną granicą wieku.

Formuła ta jest równoważna założeniu, że prawdopodobieństwo ${}_xp_0$ przeżycia przez noworodka wieku x maleje liniowo wraz ze wzrostem x , czyli

$${}_xp_0 = 1 - \frac{x}{X} \quad \text{dla} \quad 0 \leq x < X \quad \text{oraz} \quad {}_xp_0 = 0 \quad \text{dla} \quad x \geq X. \quad (4.46)$$

Kolejne dwa modele Lamberta i Gompertza pochodzą z lat odpowiednio 1776 oraz 1825. Lambert zaproponował model z czterema parametrami, który można zapisać w postaci

$${}_x p_0 = \left[\frac{a-x}{x} \right]^2 - b \left[e^{-\frac{x}{c}} - e^{-\frac{x}{a}} \right], \quad x > 0, \quad (4.47)$$

natomiast drugi model sformułowany został dla „siły umieralności”, tj. funkcji intensywności zgonów, która w modelu Gompertza rośnie wykładniczo wraz z wiekiem

$$\mu_x = Bc^x, \quad B > 0, c \geq 1, x \geq 0, \quad (4.48)$$

co z kolei prowadzi do formuły dla prawdopodobieństwa przeżycia ${}_x p_0$

$${}_x p_0 = e^{-k(c^x - 1)}, \quad (4.49)$$

gdzie $k = -B/\ln c$.

W roku 1867 Makeham zmodyfikował propozycję Gompertza wyrażoną równaniem (4.48), dodając po prawej stronie stałą A niezależną od wieku x , czyli

$$\mu_x = A + Bc^x, \quad A, B > 0, c \geq 1, x \geq 0, \quad (4.50)$$

co sprowadza się do wzoru na prawdopodobieństwo ${}_x p_0$ postaci

$${}_x p_0 = e^{-k(bx + c^x - 1)}, \quad (4.51)$$

gdzie $b = A \ln c / B$ oraz $k = -B / \ln c$.

Model(4.50) znany jest współcześnie pod nazwą prawa umieralności Gompertza–Makehama.

Inny model zaproponowany przez Weibulla w 1939 roku ma postać

$$\mu_x = kx^m, \quad k, m > 0, x > 0. \quad (4.52)$$

Intensywność zgonów w modelu Weibulla jest funkcją monotoniczną, tj. rosnącą lub malejącą, dlatego modelu tego rzadko używa się w demografii do modelowania umieralności. Częściej natomiast znajduje on zastosowanie w analizie niezawodności, np. w badaniach dotyczących czasu bezawaryjnej pracy urządzeń.

Przegląd innych parametrycznych modeli umieralności znaleźć można m.in. w pracy zbiorowej [Tabeau et al. 2001].

Ćwiczenia

Ćwiczenie 4.1

Korzystając z danych zawartych w tabelicy 10 (strona 82), obliczyć prawdopodobieństwo:

1. Dożycia wieku 25 lat, wśród kobiet w wieku 20 lat,
2. Zgonu przed upływem 5 lat, wśród kobiet w wieku 40 lat.

Ćwiczenie 4.2

Czas życia w pewnej populacji ludzi ma rozkład Gompertza z funkcją intensywności

$$\lambda(x) = \gamma \cdot e^{\delta \cdot x},$$

przy czym parametry w tym rozkładzie są równe

$$\gamma = 0,000133, \quad \delta = 0,08.$$

Na podstawie odpowiednich wzorów zbudować tabelicę trwania życia w podziale wg rocznych grup wieku, zawierającą następujące dane:

1. Liczby dożywających l_x , przyjmując liczebność początkową $l_0 = 100\,000$,
2. Liczby zgonów d_x oraz prawdopodobieństwa zgonu w ciągu roku q_x dla osób dożywających wieku x lat,
3. Fundusz osobolet T_x i średnie dalsze trwanie życia e_x dla osób dożywających wieku x lat.

Rozdział 5

REPRODUKCJA LUDNOŚCI

5.1. Wprowadzenie

Ważny dział analizy demograficznej dotyczy reprodukcji ludności, tj. zagadnień odtwarzania stanu liczebnego populacji pod wpływem ruchu naturalnego i wędrownego. W węższym sensie, reprodukcję rozważa się w odniesieniu do ruchu naturalnego, na który składają się strumienie urodzeń i zgonów, wpływające odpowiednio za przyrost i ubytek liczebny populacji.

Warto zauważyć, że na reprodukcję ludności ma także wpływ struktura ludności wg wieku. Przykładowo, w populacji z dużym udziałem osób starszych liczba urodzeń jest często niższa od liczby zgonów, co w efekcie skutkuje ujemnym przyrostem naturalnym. W tym znaczeniu struktura populacji wg wieku wpływa na potencjał odtwarzania się populacji.

Analiza urodzeń jest bardziej złożona niż analiza zgonów. W przeciwieństwie do uniwersalnego prawa umieralności, któremu podlegamy wszyscy, szansa lub swoiście rozumiane „ryzyko” urodzenia dziecka dotyczy jedynie kobiet w wieku rozrodczym. Ponadto, poza różnicami indywidualnymi, związanymi m.in. z wiekiem, zdrowiem, występuje szereg czynników społeczno-kulturowych, które mogą wykluczać w sposób czasowy lub trwały niektóre kobiety z populacji podlegającej temu „ryzyku” (np. antykoncepcja, sterylizacja itp.).

Zasadnicza różnica pomiędzy analizą urodzeń i zgonów polega jednak na tym, że wydanie na świat dziecka może być dla kobiety doświadczeniem powtarzalnym, podczas gdy zgon jest zdarzeniem jednorazowym. Oznacza to, że urodzenia można rozważać także pod kątem ich kolejności, natomiast kobiety pod kątem liczby dotychczas urodzonych dzieci (*parity*).

Analizę urodzeń odnosi się tradycyjnie do populacji kobiet. Jednak formalnie rzecz biorąc, przedstawione dalej miary reprodukcji mogą być dostosowane także do populacji mężczyzn.

5.2. Wskaźnik dynamiki demograficznej

Wskaźnik dynamiki demograficznej DDR (*demographic dynamics ratio*) jest najprostszym miernikiem reprodukcji, określającym relację liczby urodzeń U_0 i zgonów D w danej populacji i w ustalonym przedziale czasu, np. w wybranym roku. Definiujemy go wzorem

$$DDR = \frac{U_0}{D}. \quad (5.1)$$

Wartość DDR większa od 1 wskazuje na nadwyżkę liczby urodzeń nad zgonami, wartość mniejsza od 1 oznacza relację odwrotną, natomiast wartość 1 oznacza, że liczba urodzeń kompensuje liczbę zgonów.

Utrzymywanie się jednej z wymienionych relacji w dłuższym okresie wskazuje na reprodukcję odpowiednio rozszerzoną ($DDR > 1$), zawężoną ($DDR < 1$) lub prostą ($DDR = 1$).

5.3. Współczynniki reprodukcji netto i brutto

Współczynnik reprodukcji netto NRR (*net reproduction rate*) jest agregatowym miernikiem reprodukcji wywodzącym się z analizy kohortowej.

Założmy, że rozważamy pewną generację kobiet o liczebności początkowej $U_0^{(fem)}$, a także populację dziewczynek żywo urodzonych przez kobiety z tej kohorty, o łącznej liczebności $B^{(fem)}$. Nazwijmy obie zbiorowości odpowiednio pokoleniem matek i pokoleniem córek. Wówczas iloraz

$$NRR = \frac{B^{(fem)}}{U_0^{(fem)}} \quad (5.2)$$

nazwiemy kohortowym **współczynnikiem reprodukcji netto**.

Kohortowy współczynnik reprodukcji netto NRR informuje, w jakim stopniu pokolenie matek jest zastąpione przez ich córki. W populacji zamkniętej, jeśli NRR utrzymuje się na poziomie większym niż 1 dla kolejnych pokoleń kobiet (tzw. reprodukcja rozszerzona),

wówczas stan liczebny pokoleń wzrasta, natomiast jeśli NRR utrzymuje się poniżej 1 (reprodukcja zawężona), wówczas liczebność kolejnych pokoleń maleje. Jeśli wartość NRR stabilizuje się na poziomie 1 (reprodukcja prosta), wówczas populacja zmierza do stanu zwanego stacjonarnym [zob. sekcja 3.7].

Współczynnik (5.2) jest średnią liczbą żywo urodzonych córek wydanych na świat przez kobiety z danej generacji. Zauważymy, że we wzorze tym zarówno liczba córek $B^{(fem)}$, jak i liczba matek $U_0^{(fem)}$ wzięte są pod uwagę przy urodzeniu [Pressat 1966, s. 406]. Pomnożmy licznik i mianownik przez prawdopodobieństwa ${}_x p_0^{(m)}$, ${}_x p_0^{(c)}$ dożycia przez noworodka płci żeńskiej do wieku rozrodczego $x = 15$ lat, odpowiednio w pokoleniu matek i ich córek. Otrzymamy następującą zmodyfikowaną formułę kohortowego współczynnika reprodukcji netto

$$NRR^* = \frac{B^{(fem)} \cdot {}_x p_0^{(c)}}{U_0^{(fem)} \cdot {}_x p_0^{(m)}}. \quad (5.3)$$

Powyższy zapis pozwala na interpretację NRR^* , jako średniej liczby córek osiągających wiek rozrodczy, przypadającej na 1 kobietę z danej generacji, dożywającą wieku rozrodczego.

Iloraz (5.3) sprowadza się do (5.2), gdy ${}_x p_0^{(c)} = {}_x p_0^{(m)}$, czyli przy założeniu, że prawdopodobieństwa dożycia do ustalonego wieku x są jednakowe w pokoleniu matek i córek. Takie założenie jest spełnione np. w populacji ustabilizowanej. Wówczas prawdopodobieństwa w liczniku i mianowniku (5.3) upraszczają się i prowadzą do wzoru (5.2). W innych przypadkach wzory (5.2) oraz (5.3) nie są równoważne.

Poniżej przedstawimy kilka równoważnych do (5.2) formuł definiujących kohortowy współczynnik reprodukcji netto NRR .

Dostępne w statystyce publicznej dane dotyczą zwykle urodzeń ogółem. W takich przypadkach licznik ilorazu (5.2) obliczamy, mnożąc łączną liczbę urodzeń przez udział dziewczynek. Prawdopodobieństwa urodzenia dziewczynki lub chłopca są uwarunkowane biologicznie, wynosząc w przybliżeniu odpowiednio 0,485 i 0,515. Warto jednak dodać, że proporcje te mogą zostać zakłócone np. w sytuacji ingerencji człowieka w naturalny proces kształtowania się płci dziecka (przykładem mogą być aborcje dzieci określonej płci lub próby ingerencji genetycznej).

Oznaczmy symbolem c iloraz urodzeń dzieci płci żeńskiej $B^{(fem)}$ do urodzeń ogółem B , czyli

$$c = \frac{B^{(fem)}}{B}. \quad (5.4)$$

Mamy wtedy z (5.2)

$$NRR = c \cdot \frac{B}{U_0^{(fem)}}. \quad (5.5)$$

Inna reprezentacja kohortowego współczynnika NRR opiera się na relacji wiążącej go z cząstkowymi współczynnikami płodności. W przypadku rocznych grup wieku $[x, x + 1)$ relacja ta jest następująca

$$NRR = c \cdot \sum_{x=a_1}^{a_2-1} f_x \cdot \frac{K_x^{(fem)}}{U_0^{(fem)}}, \quad (5.6)$$

gdzie $[a_1, a_2)$ jest przedziałem wieku prokreacyjnego (zakładamy dalej $a_1 = 15, a_2 = 50$ lat), f_x dla $x = 15, \dots, 49$ są kohortowymi, cząstkowymi współczynnikami płodności dla rocznych grup wieku, $K_x^{(fem)}$ jest czasem ekspozycji kobiet w grupie wieku $[x, x + 1)$ lat, natomiast $U_0^{(fem)}$ oznacza liczebność początkową kohorty kobiet (pokolenia matek).

Zauważmy, że zapisy (5.5) i (5.6) są równoważne. Przypomnijmy, że cząstkowe współczynniki płodności f_x dla rocznych grup wieku definiujemy jako ilorazy $B_x/K_x^{(fem)}$, gdzie B_x oznacza liczbę dzieci urodzonych przez kobiety w wieku $[x, x + 1)$ lat, a $K_x^{(fem)}$ jest czasem ekspozycji kobiet w tej grupie wieku [por. (2.33)]. Po podstawieniu wspomnianych ilorazów w miejsce f_x po prawej stronie (5.6), formuła sprowadza się do sumy urodzeń $\sum_x B_x$ podzielonej przez $U_0^{(fem)}$. Z kolei suma $\sum_x B_x$ dla x przebiegającego przedział wieku $[a_1, a_2)$ daje łączną liczbę B dzieci urodzonych przez kobiety w wieku rozrodczym, co w efekcie prowadzi do wzoru (5.5).

Niekiedy w miejsce iloczynów $c \cdot f_x$ we wzorze (5.6) przyjmuje się cząstkowe współczynniki płodności dotyczące urodzeń żeńskich, które oznaczymy symbolem $f_x^{(fem)}$. Wtedy (5.6) przyjmuje postać

$$NRR = \sum_{x=a_1}^{a_2-1} f_x^{(fem)} \cdot \frac{K_x^{(fem)}}{U_0^{(fem)}}. \quad (5.7)$$

Gdy dysponujemy cząstkowymi współczynnikami płodności w zakresie urodzeń żeńskich dla szerszych przedziałów wieku $[x, x + n)$, tj. gdy $n > 1$ ($n \in \mathbf{N}$), wówczas NRR możemy wyznaczać ze wzoru

$$NRR = \sum_x n f_x^{(fem)} \cdot \frac{n K_x^{(fem)}}{U_0^{(fem)}}. \quad (5.8)$$

Liczby osobolat życia $K_x^{(fem)}$ w formule (5.7) lub $n K_x^{(fem)}$ w (5.8) zwykle nie są znane. Przybliżyć je możemy przy pomocy liczb kobiet dożywających wieku odpowiednio $x + \frac{1}{2}$ lub $x + \frac{n}{2}$ pomnożonych przez szerokość przedziałów wieku, tj. za pomocą wyrażeń $U_{x+\frac{1}{2}}^{(fem)}$ lub $n \cdot U_{x+\frac{n}{2}}^{(fem)}$ [przez analogię do (2.4) lub (2.8)]. Ponadto, ilorazy

$$\frac{U_{x+\frac{1}{2}}^{(fem)}}{U_0^{(fem)}}, \quad \frac{U_{x+\frac{n}{2}}^{(fem)}}{U_0^{(fem)}} \quad (5.9)$$

traktować możemy jako prawdopodobieństwa dożycia wieku $x + \frac{1}{2}$ lub $x + \frac{n}{2}$ przez noworodka płci żeńskiej. Oznaczmy te prawdopodobieństwa symbolami odpowiednio ${}_{x+\frac{1}{2}}p_0$ oraz ${}_{x+\frac{n}{2}}p_0$. Wówczas NRR dla rocznych grup wieku $[x, x + 1)$ można przybliżać wzorami

$$NRR \approx \sum_{x=a_1}^{a_2-1} f_x^{(fem)} \cdot {}_{x+\frac{1}{2}}p_0 \quad \text{lub} \quad NRR \approx c \sum_{x=a_1}^{a_2-1} f_x \cdot {}_{x+\frac{1}{2}}p_0, \quad (5.10)$$

a w przypadku grup wieku $[x, x + n)$ wzorami

$$NRR \approx n \sum_x n f_x^{(fem)} \cdot {}_{x+\frac{n}{2}}p_0 \quad \text{lub} \quad NRR \approx n c \sum_x n f_x \cdot {}_{x+\frac{n}{2}}p_0, \quad (5.11)$$

przy czym prawdopodobieństwa ${}_{x+\frac{1}{2}}p_0$ oraz ${}_{x+\frac{n}{2}}p_0$ wyznaczamy z kohortowych tablic trwania życia dla kobiet wg formuł

$${}_{x+\frac{1}{2}}p_0 = \frac{1}{l_0} \cdot \frac{l_x + l_{x+1}}{2}, \quad (5.12)$$

$${}_{x+\frac{n}{2}}p_0 = \frac{1}{l_0} \cdot \frac{l_x + l_{x+n}}{2},$$

gdzie l_x oznacza liczby kobiet dożywających wieku x .

Warto wspomnieć, że współczynnik reprodukcji netto oblicza się z reguły w ujęciu przekrojowym. NRR obliczamy wówczas ze wzorów (5.10) lub (5.11), zastępując kohortowe współczynniki płodności odpowiednimi współczynnikami przekrojowymi oraz wyznaczając prawdopodobieństwa (5.12) z przekrojowych tablic trwania życia dla kobiet.

Kolejny wskaźnik, zwany **współczynnikiem reprodukcji brutto** GRR (*gross reproduction rate*), ma charakter teoretyczny. Wyznaczamy go na podstawie cząstkowych współczynników płodności przy założeniu braku umieralności. Podobnie jak współczynnik reprodukcji netto, GRR wywodzi się z analizy kohortowej.

Założmy, że czas ekspozycji na szansę urodzenia dziecka dla każdej kobiety w danej grupie wieku $[x, x + 1)$ równy jest dokładnie długości tego przedziału, tj. wynosi 1 rok, co jest równoznaczne z założeniem braku umieralności w tym przedziale wieku. Wówczas formuła definiująca GRR sprowadza się do zastąpienia składnika $K_x^{(fem)}$ we wzorze (5.6) liczebnością $U_0^{(fem)}$. Otrzymujemy GRR postaci

$$GRR = c \cdot \sum_{x=a_1}^{a_2-1} f_x \frac{U_0^{(fem)}}{U_0^{(fem)}} = c \cdot \sum_{x=a_1}^{a_2-1} f_x, \quad (5.13)$$

gdzie c jest określone wzorem (5.4).

Dokonując analogicznej zamiany we wzorze (5.7), mamy także

$$GRR = \sum_{x=a_1}^{a_2-1} f_x \frac{U_0^{(fem)}}{U_0^{(fem)}} = \sum_{x=a_1}^{a_2-1} f_x. \quad (5.14)$$

Dla grup wieku postaci $[x, x + n)$, gdy $n > 1$, współczynnik GRR przyjmuje postać analogiczną do (5.13) lub (5.14), tj.

$$GRR = c \cdot n \cdot \sum_{x=a_1}^{a_2-1} n f_x \quad (5.15)$$

lub

$$GRR = n \cdot \sum_x n f_x^{(fem)}. \quad (5.16)$$

W przypadku danych przekrojowych współczynnik GRR możemy zdefiniować podobnie jak w (5.15) czy (5.16), z tym że kohortowe, cząstkowe współczynniki płodności zastępujemy współczynnikami przekrojowymi.

Przy obliczaniu współczynnika GRR możemy także skorzystać z relacji wiążącej go ze współczynnikiem dzietności TFR , o którym mowa w kolejnej sekcji.

5.4. Współczynnik dzietności

Ogólny współczynnik dzietności TFR (*total fertility rate*) jest proporcjonalny do współczynnika GRR . Pomiedzy GRR oraz TFR zachodzi relacja

$$TFR = \frac{1}{c} \cdot GRR. \quad (5.17)$$

Różnica polega na tym, że w TFR uwzględniamy dodatkowo urodzenia dzieci płci męskiej. Dla rocznych grup wieku otrzymujemy relację

$$TFR = \sum_{x=a_1}^{a_2-1} f_x, \quad (5.18)$$

natomiast w przypadku grup wieku $[x, x + n)$ dla $n > 1$ powyższa formuła przyjmuje postać

$$TFR = n \cdot \sum_x f_x. \quad (5.19)$$

Obydwa wskaźniki, GRR oraz TFR , mają charakter hipotetyczny, a ich interpretacja jest zbliżona. Współczynnik TFR określa, ile dzieci wydałaby na świat typowa kobieta w hipotetycznych warunkach braku umieralności i przy obserwowanym profilu płodności, reprezentowanym przez cząstkowe współczynniki płodności.

Patrząc jednak na zagadnienie z perspektywy reprodukcji, możemy powiedzieć, że kobieta reprodukuje się poprzez córki. Z tej perspektywy zastosowanie znajduje współczynnik reprodukcji brutto GRR wskazujący, ile córek wydałaby na świat typowa kobieta w warunkach braku umieralności i przy obserwowanym profilu płodności.

Zauważymy, że wartość współczynnika TFR mniejsza lub równa 2 nie gwarantuje prostej zastępowalności pokoleń. Minimalny poziom TFR wskazujący na prostą zastępowalność winien kształtować się na poziomie nieco powyżej 2. Uwzględnić należy bowiem efekt umieralności, pominięty w konstrukcji współczynnika TFR . W krajach wysoko rozwiniętych przyjmuje się, że prosta zastępowalność występuje w przypadku, gdy $TFR = 2,1$.

Przykład 16. Tablica 15 zawiera dane wejściowe i obliczenia pośrednie prowadzące do wyznaczenia współczynników TFR , GRR oraz NRR dla Polski w roku 2017.

Kolumny 2 i 3 tablicy zawierają ilorazy l_x/l_0 , zaczerpnięte z tablicy trwania życia dla kobiet (por. tablica 10, strona 82), a także wyznaczone na ich podstawie prawdopodobieństwa

$${}_{x+2,5}p_0 \approx (l_x + l_{x+5})/2l_0.$$

Liczby urodzeń ${}_5B_x$ w podziale wg wieku matek oraz liczby $L_{[x,x+5]}^{(fem)}$ kobiet będących w połowie roku 2017 w grupach wieku $[x, x + 5)$ lat dla $x = 15, 20, \dots, 45$ zaczerpnięto z bazy GUS (stat.gov.pl). Dane te zamieszczono w kolumnach 4 i 5.

Kolejne dwie kolumny tablicy 15 zawierają obliczenie pośrednie, w tym cząstkowych współczynników płodności, niezbędne do wyznaczenia współczynnika dzietności oraz współczynników reprodukcji brutto i netto.

Tablica 15. Dane wejściowe l_x/l_0 , ${}_5B_x$, $L_{[x,x+5]}^{(fem)}$ i obliczenia pośrednie współczynnika dzietności TFR i współczynników reprodukcji brutto GRR i netto NRR

wiek x	l_x/l_0	${}_{x+2,5}p_0$	${}_5B_x$	$L_{[x,x+5]}^{(fem)}$	${}_5f_x$	(3) × (6)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
15	0,99495	0,9944	10032	922568	0,0109	0,0108
20	0,99389	0,9933	55928	1104996	0,0506	0,0503
25	0,99269	0,9920	132464	1336564	0,0991	0,0983
30	0,99137	0,9904	134904	1566263	0,0861	0,0853
35	0,98946	0,9879	57563	1547565	0,0372	0,0367
40	0,98643	0,9838	10642	1421700	0,0075	0,0074
45	0,98119	0,9767	397	1197465	0,0003	0,0002
50	0,97227			Razem	0,2917	0,2891

Źródło: obliczenia własne na podstawie danych GUS, stat.gov.pl

Cząstkowe współczynniki płodności ${}_5f_x$ wyznaczone zostały poprzez podzielenie wyrazów z kolumny 4 przez wyrazy z kolumny 5. Wyniki dzielenia prezentuje kolumna 6.

Zgodnie z formułą (5.19), suma współczynników płodności ${}_5f_x$ pomnożona przez szerokość przedziałów wieku $n = 5$ daje współczynnik dzietności TFR . W tym przykładzie mamy

$$TFR = 5 \cdot 0,2917 = 1,4587.$$

Z kolei zgodnie z formułą (5.17), współczynnik TFR pomnożony przez udział noworodków płci żeńskiej c prowadzi do współczynnika reprodukcji brutto GRR . Przyjmując $c = 0,485$, otrzymujemy

$$GRR = 0,485 \cdot 1,4587 = 0,7075.$$

W ostatniej kolumnie tablicy 15 zawarto iloczyn prawdopodobieństw ${}_{x+2,5}p_0$ i współczynników płodności ${}_5f_x$. Na mocy (5.11) suma tych iloczynów pomnożona przez iloczyn $nc = 5 \cdot 0,485$ daje współczynnik reprodukcji netto NRR , czyli

$$NRR = 5 \cdot 0,485 \cdot 0,2891 = 0,701.$$

Ćwiczenie

Tablica 16 zawiera przekrojowe, cząstkowe współczynniki płodności w Polsce obliczone na podstawie danych GUS (stat.gov.pl) dla trzech lat.

Tablica 16. Cząstkowe współczynniki płodności w Polsce w wybranych latach

wiek x	wsp. płodności ${}_5f_x$		
	1980	2000	2007
15	0,033	0,017	0,015
20	0,180	0,083	0,058
25	0,136	0,095	0,092
30	0,069	0,052	0,067
35	0,029	0,021	0,026
40	0,008	0,005	0,005
45	0,001	0,000	0,000

Źródło: na podstawie danych GUS

Obliczyć i zinterpretować:

1. Współczynniki dzietności TFR osobno dla każdego roku,
2. Współczynniki reprodukcji brutto GRR dla badanych lat, przyjmując udział urodzeń żeńskich $c = 0,485$,
3. Współczynniki reprodukcji netto NRR , korzystając z ilorazów l_x/l_0 z przekrojowych TTŻ dla kobiet z odpowiednich lat (stat.gov.pl).

Uzyskane w punktach 1–3 wartości porównać z analogicznymi wynikami z przykładu 16.

Rozdział 6

WZROST LICZEBNY POPULACJI

6.1. Wprowadzenie

Istnieją dwa rodzaje zdarzeń demograficznych wpływających na wzrost liczebny populacji, tj. **urodzenia i napływ migracyjny** jednostek do badanej populacji, a także dwa rodzaje zdarzeń demograficznych, które mogą zredukować liczebność danej zbiorowości, tj. **zgony i odpływ migracyjny**.

Podstawowe równanie opisujące zmiany w czasie w zakresie stanu liczebnego populacji nosi miano **równania bilansowego** (*balancing equation*). Będzie ono punktem wyjścia do wyprowadzenia w dalszej części tego rozdziału jednego z bardziej znanych modeli wzrostu populacji, tzw. modelu wzrostu wykładniczego.

6.2. Równanie bilansowe

Oznaczmy przez $t = [\tau_1, \tau_2]$ ustalony przedział czasu kalendarzowego, gdzie $\tau_1 < \tau_2$ wyznaczają konkretne chwile czasu. Przykładowo, może to być okres od 1 stycznia do 31 grudnia wybranego roku t . Zmiany w rozmiarze liczebnym populacji w danym przedziale t można przedstawić w postaci następującej tożsamości, zwanej równaniem bilansowym

$$L(\tau_2) \equiv L(\tau_1) + B(t) - D(t) + NM(t), \quad (6.1)$$

gdzie:

$L(\tau_i)$ – liczebność populacji w chwili τ_i , $i = 1, 2$,

$B(t)$ – liczba urodzeń żywych w okresie $t = [\tau_1, \tau_2]$,

$D(t)$ – liczba zgonów w okresie $t = [\tau_1, \tau_2]$,

$NM(t)$ – saldo migracji (migracja netto) w okresie $t = [\tau_1, \tau_2]$.

Odejmujemy stronami początkową liczebność $L(\tau_1)$ badanej zbiorowości, a następnie podzielimy przez liczbę osobolat życia $K(t)$ jednostek populacji w okresie $t = [\tau_1, \tau_2]$. Wówczas otrzymamy

$$\frac{L(\tau_2) - L(\tau_1)}{K(t)} = \frac{B(t)}{K(t)} - \frac{D(t)}{K(t)} + \frac{NM(t)}{K(t)}. \quad (6.2)$$

Przyjmijmy następujące oznaczenia i nazwy poszczególnych ilorazów:
Współczynnik wzrostu (*crude rate of total increase*)

$$CRTI(t) = \frac{L(\tau_2) - L(\tau_1)}{K(t)}, \quad (6.3)$$

Ogólny współczynnik urodzeń (*crude birth rate*)

$$CBR(t) = \frac{B(t)}{K(t)}, \quad (6.4)$$

Ogólny współczynnik zgonów (*crude death rate*)

$$CDR(t) = \frac{D(t)}{K(t)}, \quad (6.5)$$

Współczynnik przyrostu naturalnego (*rate of natural increase*)

$$RNI(t) = CBR(t) - CDR(t) = \frac{B(t)}{K(t)} - \frac{D(t)}{K(t)}, \quad (6.6)$$

Współczynnik migracji netto (*net migration rate*)

$$NMR(t) = \frac{NM(t)}{K(t)}. \quad (6.7)$$

Ogólne współczynniki urodzeń, zgonów czy migracji netto zdefiniowane są tutaj jako liczby określonych zdarzeń demograficznych w danej populacji i w ustalonym okresie t , podzielone przez liczbę osobolat życia $K(t)$ jednostek tej populacji, w tym okresie.

Przy wykorzystaniu oznaczeń (6.3)–(6.7), równanie (6.2) prowadzi do następującej dekompozycji współczynnika wzrostu

$$CRTI(t) = CBR(t) - CDR(t) + NMR(t). \quad (6.8)$$

Warto zauważyć, że jeden dzień życia jednostki w określonej populacji stanowi wkład $1/365$ do liczby osobolat życia $K(t)$, natomiast każda godzina to wkład rzędu $1/8760$ osobolat. W każdym przypadku wielkość $K(t)$ jest wyrażona w osobolatach, niezależnie od tego, czy rozważany przedział czasu ogranicza się do godzin, dni, miesięcy, kwartałów czy też obejmuje kilka lat.

Na ogół dokładna liczba osobolat życia $K(t)$ nie jest znana. Za jej przybliżenie przyjmujemy wówczas średnią liczbę ludności w danym okresie lub liczbę ludności w połowie tego okresu. Liczba ta oznaczona będzie dalej symbolem $\bar{L}(t)$.

Przyjmując zatem przybliżenie $K(t) \approx \bar{L}(t)$, mamy na podstawie (6.3)–(6.7)

$$\begin{aligned} CRTI(t) &\approx \frac{L(\tau_1) - L(\tau_2)}{\bar{L}(t)}, \\ CBR(t) &\approx \frac{B(t)}{\bar{L}(t)}, \\ CDR(t) &\approx \frac{D(t)}{\bar{L}(t)}, \\ RNI(t) &\approx CBR(t) - CDR(t), \\ NMR(t) &\approx \frac{NM(t)}{\bar{L}(t)}. \end{aligned} \tag{6.9}$$

Przykład 17. Równanie bilansowe (6.1) dla populacji całego świata w okresie obejmującym lata 2010 – 2015 można zapisać następująco (dane World Population Prospects, <https://population.un.org/wpp/>):

$$7\,383\,008\,770 = 6\,958\,169\,159 + 701\,346\,033 - 276\,506\,422 + 0$$

Zauważymy, że saldo migracji dla świata wynosi 0, co w przypadku mniejszych regionów nie musi mieć miejsca.

Przykład 18. Rozważymy przyrost rzeczywisty liczby ludności Polski w roku 2017, tj. w okresie od 1.01.2017 do 31.12.2017 (stat.gov.pl).

Mamy

$$\underbrace{38\,433\,558 - 38\,432\,992}_{\text{przyrost ludności}=566} = \underbrace{401\,982 - 402\,852}_{\text{przyrost naturalny}=-870} + \underbrace{13\,324 - 11\,888}_{\text{saldo migracji}=1436}$$

Do mianownika wstawmy średnią liczbę ludności w roku 2017, tj. 38 433 275. Uzyskujemy w ten sposób następującą dekompozycję współczynnika wzrostu rzeczywistego:

$$\frac{566}{38\,433\,275} = \frac{401\,982}{38\,433\,275} - \frac{402\,852}{38\,433\,275} + \frac{1\,436}{38\,433\,275},$$

co po zaokrągleniu do 5 cyfr po przecinku daje równanie postaci:

$$0,00002 = 0,01046 - 0,01048 + 0,00004.$$

Oznacza to, że na przyrost rzeczywisty rzędu ok. 2 osób na 100 000 mieszkańców Polski w roku 2017 składa się 1046 urodzeń żywych na 100 000 ludności, pomniejszone o 1048 zgonów na 100 000 ludności i powiększone o dodatnie saldo migracji na poziomie ok. 4 osób na 100 000 ludności. Zatem stopa przyrostu naturalnego wynosi w tym przypadku $-0,00002$, natomiast stopa przyrostu rzeczywistego z uwzględnieniem migracji wynosi $0,00002$.

Przykład 19. Poniżej podane są współczynniki (6.9) dla świata oraz wybranych kontynentów w odniesieniu do okresu 2010–2015 (dane: World Population Prospects, <https://population.un.org/wpp/>), w układzie analogicznym do (6.8).

Świat

$$0,0119 = 0,0196 - 0,0077 + 0$$

Afryka

$$0,0259 = 0,0359 - 0,0094 - 0,0006$$

Azja

$$0,0104 = 0,0177 - 0,0070 - 0,0003$$

Europa

$$0,0010 = 0,0108 - 0,0109 + 0,0011$$

Ameryka Południowa i Środkowa

$$0,0114 = 0,0178 - 0,0058 - 0,0006$$

Ameryka Północna

$$0,0075 = 0,0124 - 0,0081 + 0,0032$$

Australia

$$0,0147 = 0,0133 - 0,0066 + 0,0080$$

Najniższy współczynnik wzrostu w latach 2010–2015 zanotowano w Europie (przyrost rządu 10 osób na 10 000 ludności). Jednak obecnie obserwowane ruchy migracyjne oraz prognozy ludnościowe dla Europy na lata 2015 – 2020 wskazują na wzrost współczynnika migracji netto w tym okresie, przy jednoczesnym spadku ogólnego współczynnika urodzeń i wzroście ogólnego współczynnika zgonów.

6.3. Intensywność wzrostu populacji

Założmy, że rozważanym przedziałem czasu $t = [\tau_1, \tau_2]$ jest ustalony rok kalendarzowy. Niech $L(1.01.t), L(2.01.t), \dots, L(31.12.t)$ oznaczają liczebności populacji w kolejnych dniach tego roku.

Za łączną liczbę osobołat życia $K(t)$ możemy przyjąć w przybliżeniu następującą sumę

$$\begin{aligned} K(t) &\approx L(1.01.t) \cdot \frac{1}{365} + L(2.01.t) \cdot \frac{1}{365} + \dots + L(31.12.t) \cdot \frac{1}{365} = \\ &= \Delta \sum_{i=1}^{365} L_i, \end{aligned} \tag{6.10}$$

gdzie L_i oznacza liczebność ludności w i -tym dniu, natomiast $\Delta = \frac{1}{365}$.

Uogólniając te rozważania, przyjmijmy, że znamy liczebność $L(\tau)$ populacji w dowolnej chwili $\tau \in [\tau_1, \tau_2]$. Założmy dodatkowo, że funkcja $L(\tau)$ jest różniczkowalna. Zatem z definicji pochodnej funkcji mamy

$$L'(\tau) = \lim_{\Delta_\tau \rightarrow 0} \frac{L(\tau + \Delta_\tau) - L(\tau)}{\Delta_\tau}, \tag{6.11}$$

z kolei liczbę osobołat życia $K(t)$ w przedziale $t = [\tau_1, \tau_2]$ możemy zapisać przez analogię do (6.10) jako

$$K(t) = \int_{\tau_1}^{\tau_2} L(\tau) d\tau. \tag{6.12}$$

Zbadamy granicę, do której dąży współczynnik wzrostu dany wzorem (6.3), gdy rozpiętość przedziału $[\tau_1, \tau_2]$ dąży do 0. W tym celu rozważać będziemy dalej przedział t oznaczony jako $t = [\tau, \tau + \Delta_\tau]$, natomiast $K(t)$ oznaczymy symbolem $K([\tau, \tau + \Delta_\tau])$. Zakładając, że Δ_τ jest małą liczbą, mamy w przybliżeniu

$$K([\tau, \tau + \Delta_\tau]) \approx L(\tau) \cdot \Delta_\tau, \quad (6.13)$$

a przybliżenie to jest tym lepsze, im bardziej Δ_τ zbliża się do 0.

Zatem korzystając z definicji współczynnika wzrostu (6.3), rozważać będziemy granicę postaci

$$\lim_{\Delta_\tau \rightarrow 0} \frac{L(\tau + \Delta_\tau) - L(\tau)}{K([\tau, \tau + \Delta_\tau])}, \quad (6.14)$$

którą oznaczymy symbolem $r(\tau)$ i nazwiemy **intensywnością wzrostu**. Pojęcie to traktujemy jako swego rodzaju analogon intensywności oprocentowania w analizach finansowych [Skalba 2002, s. 18].

Mamy

$$\begin{aligned} r(\tau) &= \lim_{\Delta_\tau \rightarrow 0} \frac{L(\tau + \Delta_\tau) - L(\tau)}{K([\tau, \tau + \Delta_\tau])} = \\ &= \lim_{\Delta_\tau \rightarrow 0} \frac{L(\tau + \Delta_\tau) - L(\tau)}{L(\tau) \cdot \Delta_\tau} = \\ &= \frac{1}{L(\tau)} \lim_{\Delta_\tau \rightarrow 0} \frac{L(\tau + \Delta_\tau) - L(\tau)}{\Delta_\tau} = \\ &= \frac{L'(\tau)}{L(\tau)} = \frac{d \ln L(\tau)}{d\tau}. \end{aligned} \quad (6.15)$$

Ostatecznie otrzymujemy

$$r(\tau) = \frac{d \ln L(\tau)}{d\tau}. \quad (6.16)$$

Całkując stronami w przedziale $[\tau_1, \tau_2]$, dostajemy

$$\int_{\tau_1}^{\tau_2} r(\tau) d\tau = \ln L(\tau) \Big|_{\tau_1}^{\tau_2} = \ln L(\tau_2) - \ln L(\tau_1) = \ln \left(\frac{L(\tau_2)}{L(\tau_1)} \right). \quad (6.17)$$

Uzyskaliśmy formułę

$$\int_{\tau_1}^{\tau_2} r(\tau) d\tau = \ln \left(\frac{L(\tau_2)}{L(\tau_1)} \right) \quad (6.18)$$

lub równoważnie

$$e^{\int_{\tau_1}^{\tau_2} r(\tau) d\tau} = \frac{L(\tau_2)}{L(\tau_1)}, \quad (6.19)$$

stąd

$$L(\tau_2) = L(\tau_1) e^{\int_{\tau_1}^{\tau_2} r(\tau) d\tau}. \quad (6.20)$$

Oznacza to, że liczebność populacji w chwili τ_2 równa jest liczebności w chwili τ_1 pomnożonej przez dodatni składnik, będący wykładniczą funkcją całki z intensywności wzrostu $r(\tau)$ w przedziale $[\tau_1, \tau_2]$.

6.3.1. Roczna stopa wzrostu – przypadki szczególne

Załóżmy, że intensywność wzrostu populacji $r(\tau)$ jest stała. Oznaczmy ją symbolem r i nazwiemy **stałą stopą wzrostu** populacji. Mamy wtedy

$$\int_{\tau_1}^{\tau_2} r(\tau) d\tau = \int_{\tau_1}^{\tau_2} r d\tau = r \int_{\tau_1}^{\tau_2} d\tau = r(\tau_2 - \tau_1), \quad (6.21)$$

stąd

$$L(\tau_2) = L(\tau_1) e^{r(\tau_2 - \tau_1)} \quad (6.22)$$

lub przyjmując $t = \tau_2 - \tau_1$, mamy

$$L(\tau_2) = L(\tau_1) e^{rt}, \quad (6.23)$$

zatem liczba ludności $L(\tau_2)$ w momencie τ_2 jest określona przez liczbę początkową $L(\tau_1)$ oraz mnożnik e^{rt} , gdzie t jest długością okresu czasu pomiędzy momentami τ_1 oraz τ_2 . Model ten nazywa się też **modelem wzrostu Malthusa**.

Zauważymy, że jeśli $r > 0$, wówczas $e^{rt} > 1$, czyli liczba ludności wzrasta. Jeśli $r < 0$, wówczas $e^{rt} < 1$, co oznacza, że liczebność populacji maleje. W przypadku szczególnym, gdy $r = 0$, tj. $e^{rt} = 1$, liczebność nie zmienia się. Utrzymywanie się tego stanu przez długi okres czasu prowadzi do modelu ludności zastojowej (stacjonarnej).

Z (6.22) mamy

$$r(\tau_2 - \tau_1) = \ln \left(\frac{L(\tau_2)}{L(\tau_1)} \right) \quad (6.24)$$

oraz

$$r = \frac{\ln \left(\frac{L(\tau_2)}{L(\tau_1)} \right)}{\tau_2 - \tau_1}. \quad (6.25)$$

Ostatnia formuła jest wygodna m.in. w obliczaniu, ile czasu musi upłynąć, aby badana populacja powiększyła się (względnie zmniejszyła) k -krotnie.

Przykład 20. Obliczymy, po jakim czasie populacja zwiększy się dwukrotnie, jeśli charakteryzuje ją stała, dodatnia stopa wzrostu:

a) $r = 0,03$ b) $r = 0,01$.

Skoro liczebność populacji ma się podwoić, mamy

$$\frac{L(\tau_2)}{L(\tau_1)} = 2.$$

Ponieważ r wyraża się formułą (6.25), otrzymujemy równanie

$$0,03 = \frac{\ln 2}{\tau_2 - \tau_1}.$$

Mamy $\ln 2 = 0,693$, więc wynika z tego dalej równanie postaci

$$0,03 = \frac{0,693}{\tau_2 - \tau_1}.$$

W rezultacie wnioskujemy, że populacja podwoi się w ciągu ok. 23 lat, ponieważ

$$\tau_2 - \tau_1 = \frac{0,693}{0,03} = 23,1.$$

Na podstawie analogicznych obliczeń można pokazać, że dla $r = 0,01$ podwojenie populacji nastąpi w ciągu 69,3 lat.

Poznaliśmy ogólną formułę (6.16) na intensywność $r(\tau)$ wzrostu populacji oraz formułę (6.20) obliczania liczby ludności w ustalonym momencie τ_2 na podstawie liczby ludności w chwili τ_1 oraz na podstawie całki z funkcji $r(\tau)$ dla $\tau \in [\tau_1, \tau_2]$. Zauważymy, że nie jest przy tym konieczna dokładna znajomość przebiegu funkcji $r(\tau)$ w tym przedziale czasowym. Wystarczy znać jej średnią wartość, ponieważ w (6.20) występuje całka oznaczona z intensywności $r(\tau)$.

Niech $\bar{r}_{[\tau_1, \tau_2]}$ oznacza dalej średnią funkcji $r(\tau)$ dla $\tau \in [\tau_1, \tau_2]$. Nazwiemy tę wielkość **średnią roczną stopą wzrostu** w przedziale czasu $[\tau_1, \tau_2]$. Innymi słowy, niech

$$\bar{r}_{[\tau_1, \tau_2]} = \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} r(\tau) d\tau. \quad (6.26)$$

Zauważymy przy tej okazji, że średnia roczna stopa wzrostu nie jest tożsama z intensywnością wzrostu populacji. Związek pomiędzy oboma wielkościami pokazuje formuła (6.26).

Korzystając z (6.26), liczebność (6.20) możemy zapisać

$$L(\tau_2) = L(\tau_1) e^{(\tau_2 - \tau_1) \bar{r}_{[\tau_1, \tau_2]}}. \quad (6.27)$$

W ten sposób otrzymujemy formułę na średnią roczną stopę wzrostu populacji w okresie $[\tau_1, \tau_2]$

$$\bar{r}_{[\tau_1, \tau_2]} = \frac{\ln \left(\frac{L(\tau_2)}{L(\tau_1)} \right)}{\tau_2 - \tau_1}. \quad (6.28)$$

Jest oczywiste, że w przypadku stałej stopy wzrostu zachodzi równość $\bar{r}_{[\tau_1, \tau_2]} = r$.

Warto także zauważyć, że średnia stopa $\bar{r}_{[\tau_1, \tau_2]}$ w okresie $t = [\tau_1, \tau_2]$ równa się w przybliżeniu współczynnikowi wzrostu $CRTI(t)$, wyrażonemu wzorem (6.3), o ile rozważany okres jest relatywnie krótki.

Pokażemy, że średnia arytmetyczna z dwóch średnich stóp $\bar{r}_{[\tau_1, \tau_2]}$ oraz $\bar{r}_{[\tau_2, \tau_3]}$ określonych na sąsiadujących przedziałach $[\tau_1, \tau_2]$, $[\tau_2, \tau_3]$ o jednakowej rozpiętości prowadzi do średniej stopy wzrostu $\bar{r}_{[\tau_1, \tau_3]}$ na przedziale łącznym $[\tau_1, \tau_3]$.

Oznaczmy w tym celu rozpiętość obu przedziałów $[\tau_1, \tau_2]$ oraz $[\tau_2, \tau_3]$ symbolem t , czyli niech $t = \tau_2 - \tau_1 = \tau_3 - \tau_2$. Mamy wtedy

$$\begin{aligned} \frac{\bar{r}_{[\tau_1, \tau_2]} + \bar{r}_{[\tau_2, \tau_3]}}{2} &= \frac{1}{2} \left[\frac{\ln \left(\frac{L(\tau_2)}{L(\tau_1)} \right)}{t} + \frac{\ln \left(\frac{L(\tau_3)}{L(\tau_2)} \right)}{t} \right] = \\ &= \frac{1}{2t} \left[\ln \left(\frac{L(\tau_2)}{L(\tau_1)} \right) + \ln \left(\frac{L(\tau_3)}{L(\tau_2)} \right) \right] = \\ &= \frac{1}{2t} \ln \left(\frac{L(\tau_2)L(\tau_3)}{L(\tau_1)L(\tau_2)} \right) = \frac{\ln \left(\frac{L(\tau_3)}{L(\tau_1)} \right)}{2t} = \bar{r}_{[\tau_1, \tau_3]}. \end{aligned} \quad (6.29)$$

Otrzymany wynik można uogólnić na przypadek średniej arytmetycznej z kilku średnich stóp określonych na sąsiadujących przedziałach o jednakowej rozpiętości.

Przykład 21. Rozważmy ludność Polski w okresie od $\tau_1 = 1.01.2016$ do $\tau_2 = 31.12.2016$. Mamy na podstawie danych GUS (stat.gov.pl)

$$L(\tau_1) = 38\,437\,239, \quad L(\tau_2) = 38\,432\,992.$$

Obliczmy średnią stopę wzrostu populacji na przestrzeni roku 2016. Długość rozważanego przedziału mierzona w latach wynosi 1, czyli

$$\tau_2 - \tau_1 = 1,$$

stad

$$\bar{r}_{2016} = \ln \left(\frac{38\,432\,992}{38\,437\,239} \right) \approx -0,00011.$$

Jednocześnie współczynnik wzrostu $CRTI$ obliczony ze wzoru (6.3) dla roku 2016 jest równy

$$CRTI(2016) = \frac{38\,432\,992 - 38\,437\,239}{38\,435\,115,5} \approx -0,00011,$$

gdzie liczba 38 435 115,5 w mianowniku jest średnia z liczb $L(\tau_1) = 38\,437\,239$ oraz $L(\tau_2) = 38\,432\,992$.

Otrzymana wartość $-0,00011$ oznacza, że w roku 2016 miał miejsce spadek liczby ludności o ok. 11 osób w przeliczeniu na 100 000 mieszkańców.

Analogiczne obliczenia wykonamy dla roku 2017, tj. dla okresu od $\tau_2 = 1.01.2017$ do $\tau_3 = 31.12.2017$. Mamy na podstawie danych GUS

$$L(\tau_2) = 38\,432\,992, \quad L(\tau_3) = 38\,433\,558,$$

stad średnia stopa wzrostu populacji wynosi

$$\bar{r}_{2017} = \ln \left(\frac{38\,433\,558}{38\,432\,992} \right) \approx 0,00001.$$

Zbliżoną wartość ma współczynnik wzrostu $CRTI(2017)$

$$CRTI(2017) = \frac{38\,433\,558 - 38\,432\,992}{38\,433\,275} \approx 0,00001.$$

Uzyskany wynik 0,00001 oznacza wzrost liczby ludności w roku 2017 o ok. 1 osobę w przeliczeniu na 100 000 mieszkańców.

Obliczymy na koniec średnią stopę wzrostu oraz współczynnik wzrostu, korzystając z danych dla okresu 2-letniego, tj. od $\tau_1 = 1.01.2016$ do $\tau_3 = 31.12.2017$. Rozpiętość rozważanego przedziału mierzona w latach wynosi $\tau_3 - \tau_1 = 2$, zatem mamy

$$\bar{r}_{[2016,2017]} = \frac{\ln\left(\frac{38\ 433\ 558}{38\ 437\ 239}\right)}{2} \approx -0,00005.$$

Jednocześnie *CRTI* dla dwuletniego okresu 2016 – 2017 wynosi w przybliżeniu

$$CRTI([2016, 2017]) = \frac{38\ 433\ 558 - 38\ 437\ 239}{2 \cdot 38\ 435\ 398,5} \approx -0,00005.$$

Otrzymaliśmy wartość ujemną oznaczającą spadek liczby populacji w latach 2016 – 2017 średnio rocznie o 5 osób w przeliczeniu na 100 000 ludności.

Ten sam wynik uzyskamy, uśredniając stopy \bar{r}_{2016} oraz \bar{r}_{2017} , czyli

$$\bar{r}_{[2016,2017]} = \frac{-0,00011 + 0,00001}{2} \approx -0,00005.$$

6.3.2. Liczba osobolat życia

Warto zauważyć, że na mocy (6.12), możemy znaleźć liczbę $K(t)$ osobolat życia jednostek populacji w zadanym przedziale czasu $t = [\tau_1, \tau_2]$. Oznaczmy dla uproszczenia przez \bar{r} średnią stopę wzrostu (6.26) w danym przedziale. Wówczas z (6.12) oraz (6.26) mamy

$$\begin{aligned} K(t) &= \int_{\tau_1}^{\tau_2} L(\tau) d\tau = L(\tau_1) \int_{\tau_1}^{\tau_2} e^{\bar{r}(\tau-\tau_1)} d\tau = \\ &= \frac{L(\tau_1)}{\bar{r}} e^{\bar{r}(\tau-\tau_1)} \Big|_{\tau_1}^{\tau_2} = \frac{L(\tau_1)}{\bar{r}} [e^{\bar{r}(\tau_2-\tau_1)} - 1] = \\ &= \frac{L(\tau_1) e^{\bar{r}(\tau_2-\tau_1)} - L(\tau_1)}{\bar{r}} = \frac{L(\tau_2) - L(\tau_1)}{\bar{r}}. \end{aligned} \quad (6.30)$$

Z drugiej strony na podstawie (6.28) mamy także

$$\bar{r} = \frac{\ln\left(\frac{L(\tau_2)}{L(\tau_1)}\right)}{\tau_2 - \tau_1}. \quad (6.31)$$

Zakładając, że $L(\tau_1) \neq L(\tau_2)$, otrzymujemy wzór na liczbę osobolat życia (czasu ekspozycji) w okresie $t = [\tau_1, \tau_2]$ postaci

$$K(t) = (\tau_2 - \tau_1) \cdot \frac{L(\tau_2) - L(\tau_1)}{\ln\left(\frac{L(\tau_2)}{L(\tau_1)}\right)}. \quad (6.32)$$

Gdy $\bar{r} = 0$, wówczas liczebność populacji $L(\tau)$ jest stała dla każdego τ , co oznacza, że

$$K(t) = (\tau_2 - \tau_1) \cdot L(\tau_1) = (\tau_2 - \tau_1) \cdot L(\tau_2). \quad (6.33)$$

Podsumowując, mamy dla $t = [\tau_1, \tau_2]$

$$K(t) = \begin{cases} (\tau_2 - \tau_1) \cdot \frac{L(\tau_2) - L(\tau_1)}{\ln\left(\frac{L(\tau_2)}{L(\tau_1)}\right)}, & \text{dla } \bar{r} \neq 0, \\ (\tau_2 - \tau_1) \cdot L(\tau_1), & \text{dla } \bar{r} = 0. \end{cases} \quad (6.34)$$

6.4. Intensywność urodzeń, płodności, zgonów i migracji

Oznaczmy liczbę urodzeń, zgonów i migracji netto w przedziale czasu $[\tau, \tau + \Delta_\tau)$ symbolami odpowiednio $B([\tau, \tau + \Delta_\tau))$, $D([\tau, \tau + \Delta_\tau))$, $NM([\tau, \tau + \Delta_\tau))$. Niech ponadto

$$\begin{aligned} b(\tau) &= \lim_{\Delta_\tau \rightarrow 0} \frac{B([\tau, \tau + \Delta_\tau))}{\Delta_\tau}, \\ d(\tau) &= \lim_{\Delta_\tau \rightarrow 0} \frac{D([\tau, \tau + \Delta_\tau))}{\Delta_\tau}, \\ nm(\tau) &= \lim_{\Delta_\tau \rightarrow 0} \frac{NM([\tau, \tau + \Delta_\tau))}{\Delta_\tau}. \end{aligned} \quad (6.35)$$

Na mocy (6.35) wyrażenia $b(\tau)d\tau$, $d(\tau)d\tau$ $nm(\tau)d\tau$ interpretujemy odpowiednio jako liczby urodzeń, zgonów i migracji netto w bardzo małym przedziale czasu $[\tau, \tau + d\tau)$.

Niech $K([\tau, \tau + \Delta_\tau))$ oznacza liczbę osobolat życia jednostek populacji w przedziale czasu $[\tau, \tau + \Delta_\tau)$, natomiast $K^{(fem)}([\tau, \tau + \Delta_\tau))$

– czas ekspozycji kobiet na szansę urodzenia dziecka w tym przedziale, lub innymi słowy – liczbę osobolat życia kobiet w wieku rozrodczym w okresie $[\tau, \tau + \Delta_\tau)$.

Zgodnie z rozważaniami przedstawionymi na początku sekcji 6.3, wielkość $K([\tau, \tau + \Delta_\tau))$ możemy przybliżyć za pomocą iloczynu

$$K([\tau, \tau + \Delta_\tau)) \approx L(\tau) \cdot \Delta_\tau, \quad (6.36)$$

podobnie mamy także

$$K^{(fem)}([\tau, \tau + \Delta_\tau)) \approx L^{(fem)}(\tau) \cdot \Delta_\tau, \quad (6.37)$$

gdzie $L(\tau)$, $L^{(fem)}(\tau)$ oznaczają odpowiednio liczbę ludności w chwili τ oraz liczbę kobiet w wieku rozrodczym w chwili τ .

Rozważmy granicę postaci

$$\begin{aligned} \beta(\tau) &= \lim_{\Delta_\tau \rightarrow 0} \frac{B([\tau, \tau + \Delta_\tau))}{K([\tau, \tau + \Delta_\tau))} = \\ &= \frac{1}{L(\tau)} \lim_{\Delta_\tau \rightarrow 0} \frac{B([\tau, \tau + \Delta_\tau))}{\Delta_\tau} = \frac{b(\tau)}{L(\tau)}. \end{aligned} \quad (6.38)$$

Analogicznie niech

$$\begin{aligned} \phi(\tau) &= \lim_{\Delta_\tau \rightarrow 0} \frac{B([\tau, \tau + \Delta_\tau))}{K^{(fem)}([\tau, \tau + \Delta_\tau))} = \\ &= \frac{1}{L^{(fem)}(\tau)} \lim_{\Delta_\tau \rightarrow 0} \frac{B([\tau, \tau + \Delta_\tau))}{\Delta_\tau} = \frac{b(\tau)}{L^{(fem)}(\tau)} \end{aligned} \quad (6.39)$$

oraz

$$\begin{aligned} \mu(\tau) &= \lim_{\Delta_\tau \rightarrow 0} \frac{D([\tau, \tau + \Delta_\tau))}{K([\tau, \tau + \Delta_\tau))} = \\ &= \frac{1}{L(\tau)} \lim_{\Delta_\tau \rightarrow 0} \frac{D([\tau, \tau + \Delta_\tau))}{\Delta_\tau} = \frac{d(\tau)}{L(\tau)}. \end{aligned} \quad (6.40)$$

$$\begin{aligned} \eta(\tau) &= \lim_{\Delta_\tau \rightarrow 0} \frac{NM([\tau, \tau + \Delta_\tau))}{K([\tau, \tau + \Delta_\tau))} = \\ &= \frac{1}{L(\tau)} \lim_{\Delta_\tau \rightarrow 0} \frac{NM([\tau, \tau + \Delta_\tau))}{\Delta_\tau} = \frac{nm(\tau)}{L(\tau)}. \end{aligned}$$

Funkcje $\beta(\tau)$, $\phi(\tau)$, $\mu(\tau)$, $\eta(\tau)$ będziemy nazywać odpowiednio **intensywnością urodzeń, płodności, zgonów i migracji netto**.

W analogiczny sposób zdefiniować można intensywność urodzeń, płodności, zgonów czy migracji netto w ustalonej chwili τ i dla ustalonego wieku x . W szczególności, rozważać będziemy dalej intensywność płodności $\phi_x(\tau)$ oraz intensywność zgonów $\mu_x(\tau)$ zdefiniowane, podobnie jak $\phi(\tau)$ i $\mu(\tau)$, w postaci następujących ilorazów

$$\phi_x(\tau) = \frac{b_x(\tau)}{L_x^{(fem)}(\tau)}, \quad (6.41)$$

$$\mu_x(\tau) = \frac{d_x(\tau)}{L_x(\tau)}, \quad (6.42)$$

gdzie $b_x(\tau)d\tau$ jest liczbą urodzeń w grupie kobiet w wieku x w okresie $[\tau, \tau + d\tau)$, $d_x(\tau)d\tau$ jest liczbą zgonów w grupie wieku x w przedziale $[\tau, \tau + d\tau)$, $L_x(\tau)$ oznacza liczbę ludności w wieku x w chwili τ , natomiast $L_x^{(fem)}(\tau)$ – liczbę kobiet w wieku x w chwili τ .

Gdy intensywność zgonów $\mu_x(\tau)$ i płodności $\phi_x(\tau)$ są stałe w czasie, wówczas będziemy pisać

$$\mu_x(\tau) \equiv \mu_x, \quad \phi_x(\tau) \equiv \phi_x. \quad (6.43)$$

Na koniec rozważań w tym paragrafie warto dodać, że możemy dokonać analogicznej do (6.8) dekompozycji intensywności wzrostu populacji $r(\tau)$ w postaci równania

$$r(\tau) = \beta(\tau) - \mu(\tau) + \eta(\tau). \quad (6.44)$$

Dalej rozważymy szczególny przypadek populacji zwanej modelem populacji ustabilizowanej, w którym zakładamy m.in. brak migracji (tzw. populacja zamknięta). W modelu tym lewa i prawa strona (6.44) redukują się do pewnych wielkości stałych, a ponadto $\eta(\tau) = 0$.

6.5. Model Lotki ludności ustabilizowanej

Matematyczny model ludności wyrażony w postaci pewnego równania całkowego przedstawili w 1911 roku Sharpe i Lotka. Rozważany w następnych pracach model ludności ustabilizowanej znany jest obecnie pod nazwą **modelu Lotki** [Lotka 1939].

Przyjmijmy następujące założenia:

- 1) populacja jest zamknięta (brak migracji),
- 2) intensywność zgonów $\mu_x(\tau)$ zależy od wieku x , lecz nie zależy od czasu kalendarzowego, czyli zapis $\mu_x(\tau)$ możemy uprościć do μ_x ,
- 3) intensywność płodności $\phi_x(\tau)$ zależy od wieku x i jednocześnie nie zależy od czasu kalendarzowego, czyli $\phi_x(\tau) \equiv \phi_x$.

Niech $0 < \tau_0 < \tau_1 < \tau_2 < \infty$ będą ustalonymi momentami czasu. Niech ponadto $b(\tau)d\tau$ oznacza liczbę dzieci urodzonych w przedziale czasu $[\tau, \tau + d\tau]$ (w oryginalnym modelu Lotki rozważania odnoszą się do urodzeń jednej płci, tj. płci żeńskiej). Wówczas całka

$$\int_{\tau_0}^{\tau_2} b(\tau)d\tau \quad (6.45)$$

oznacza liczbę urodzeń w okresie $[\tau_0, \tau_2]$.

Załóżmy, że liczba urodzeń $b(\tau)d\tau$ zmienia się w czasie wg wzoru

$$b(\tau) = b(\tau_0)e^{\rho(\tau-\tau_0)}, \quad \tau > \tau_0, \quad (6.46)$$

Równoważnie (6.46) możemy zapisać w postaci

$$b(\tau_0) = b(\tau)e^{-\rho(\tau-\tau_0)} \quad \tau_0 < \tau. \quad (6.47)$$

Parametr ρ nazwiemy roboczo stopą wzrostu liczby urodzeń. Pokażemy, że przy założeniach 1–3 parametr ρ jest równy stopie wzrostu populacji wyrażonej formułą analogiczną do (6.25). W tym celu oznaczymy $\tau_0 = \tau_1 - x$. W momencie $\tau_0 = \tau_1 - x$ na mocy (6.47) mamy

$$b(\tau_0) = b(\tau_1)e^{-\rho x}, \quad (6.48)$$

natomiast w momencie τ_2 na mocy (6.46) mamy

$$b(\tau_2) = b(\tau_1)e^{\rho(\tau_2-\tau_1)}. \quad (6.49)$$

Niech $S(x) = {}_x p_0$ oznacza prawdopodobieństwo dożycia noworodka do wieku x . Prawdopodobieństwo to nie zależy od τ , zgodnie z przyjętym założeniem o niezmienniej w czasie intensywności zgonów μ_x [por. formuła (4.14)]. Mnożąc stronami (6.48) przez $S(x)dx$ oraz (6.49) przez $S(x)e^{-\rho x}dx$, otrzymujemy odpowiednio

$$S(x)b(\tau_0)dx = S(x)b(\tau_1)e^{-\rho x}dx, \quad (6.50)$$

$$S(x)b(\tau_2)e^{-\rho x}dx = S(x)b(\tau_1)e^{\rho(\tau_2-\tau_1)}e^{-\rho x}dx. \quad (6.51)$$

Wyrażenie $S(x)b(\tau_0)dx$ możemy interpretować jako liczbę osób będących w chwili τ_1 w grupie wieku $[x, x+dx)$. Podobnie $S(x)b(\tau_2)e^{-\rho x}dx$ interpretujemy jako liczbę osób będących w chwili τ_2 w przedziale wieku $[x, x+dx)$.

Łączna liczebność populacji $L(\tau_1)$ w chwili τ_1 we wszystkich grupach wieku jest więc równa

$$\begin{aligned} L(\tau_1) &= \int_0^\infty S(x)b(\tau_0)dx = \int_0^\infty S(x)b(\tau_1)e^{-\rho x}dx = \\ &= b(\tau_1) \int_0^\infty S(x)e^{-\rho x}dx. \end{aligned} \quad (6.52)$$

Podobnie, łączna liczebność populacji w chwili τ_2 we wszystkich grupach wieku wynosi

$$\begin{aligned} L(\tau_2) &= \int_0^\infty S(x)b(\tau_2)e^{-\rho x}dx = \\ &= \int_0^\infty S(x)b(\tau_1)e^{\rho(\tau_2-\tau_1)}e^{-\rho x}dx = \\ &= e^{\rho(\tau_2-\tau_1)}b(\tau_1) \int_0^\infty S(x)e^{-\rho x}dx. \end{aligned} \quad (6.53)$$

Otrzymaliśmy wyrażenie analogiczne do (6.22), czyli

$$L(\tau_2) = L(\tau_1)e^{\rho(\tau_2-\tau_1)}, \quad (6.54)$$

co z kolei prowadzi do formuły zgodnej z (6.25)

$$\rho = \frac{\ln\left(\frac{L(\tau_2)}{L(\tau_1)}\right)}{\tau_2 - \tau_1}. \quad (6.55)$$

Wynika z powyższego, że zachodzi równość

$$r = \rho, \quad (6.56)$$

a zatem przy założeniach 1–3 stopa wzrostu populacji jest równa stopie wzrostu liczby urodzeń. Z tego powodu stopę ρ określa się zamiennie mianem **rzeczywistego współczynnika przyrostu naturalnego**, **współczynnika Lotki** lub **istotnej stopy wzrostu** (*intrinsic rate of growth*).

6.5.1. Struktura ludności w modelu Lotki

Rozważmy udział $w_x dx$ osób będących w wieku $[x, x + dx)$ w populacji ogółem. W przypadku ludności ustabilizowanej udział ten jest stały w czasie, tj. nie zależy od τ . Mamy bowiem z (6.50) oraz (6.52) dla $\tau = \tau_1$

$$w_x dx = \frac{S(x)b(\tau_1)e^{-\rho x} dx}{b(\tau_1) \int_0^\infty S(x)e^{-\rho x} dx} = \frac{S(x)e^{-\rho x} dx}{\int_0^\infty S(x)e^{-\rho x} dx}. \quad (6.57)$$

Podobny iloraz otrzymamy dla $\tau = \tau_2$ na podstawie (6.51) oraz (6.54)

$$w_x dx = \frac{S(x)b(\tau_1)e^{\rho(\tau_2-\tau_1)}e^{-\rho x} dx}{e^{\rho(\tau_2-\tau_1)}b(\tau_1) \int_0^\infty S(x)e^{-\rho x} dx} = \frac{S(x)e^{-\rho x} dx}{\int_0^\infty S(x)e^{-\rho x} dx}. \quad (6.58)$$

Co więcej, analogiczny iloraz otrzymamy dla dowolnego τ .

Oznacza to, że choć liczebności bezwzględne poszczególnych grup wieku zmieniają się w czasie, np. rosną, gdy $\rho > 0$ lub maleją, gdy $\rho < 0$, to jednak struktura populacji wg wieku pozostaje niezmienna. Własność powyższa wyjaśnia nazwę modelu. Szczególnym przypadkiem modelu ludności ustabilizowanej jest model ludności stacjonarnej (zastojowej), w którym $\rho = 0$.

Z modelu Lotki płynie zatem ważny wniosek. Stabilność w czasie poziomów umieralności i płodności w populacji zamkniętej prowadzi do ukształtowania się ludności o stałej strukturze według wieku i stałym przyroście lub ubytku ludności reprezentowanym przez rzeczywisty współczynnik przyrostu naturalnego ρ (dodatni lub ujemny).

Współczynnik ρ jest rozwiązaniem tzw. integralnego równania reprodukcji ludności, przy zadanych, stałych poziomach płodności i umieralności. Równanie to wyprowadzone zostało w kolejnej sekcji.

6.5.2. Integralne równanie reprodukcji ludności

Niech $b^{(fem)}(\tau)d\tau$ oznacza liczbę noworodków płci żeńskiej urodzonych w przedziale czasu $[\tau, \tau + d\tau)$. Niech ponadto ϕ_x oznacza stałą w czasie intensywność płodności kobiet w wieku x .

Wówczas, przez analogię do (6.50), liczba kobiet będących w chwili τ w wieku $[x, x + dx)$ jest równa

$$S^{(fem)}(x)b^{(fem)}(\tau)e^{-\rho x}dx, \quad (6.59)$$

gdzie $S^{(fem)}(x)$ oznacza prawdopodobieństwo dożycia noworodka płci żeńskiej do wieku x .

Z kolei liczba dziewczynek urodzonych w chwili τ przez kobiety w wieku $[x, x + dx)$, czyli wyrażenie $b_x^{(fem)}(\tau)dx$ wynosi

$$S^{(fem)}(x)b^{(fem)}(\tau)\phi_x e^{-\rho x}dx. \quad (6.60)$$

Całka oznaczona wyrażenia (6.60) dla $x \in [a_1, a_2)$ wynosi $b^{(fem)}(\tau)$. Tym samym iloczyn $b^{(fem)}(\tau)d\tau$, reprezentujący łączną liczbę dziewczynek urodzonych w przedziale czasu $[\tau, \tau + d\tau)$ przez wszystkie kobiety w wieku $[a_1, a_2)$, wynosi

$$b^{(fem)}(\tau)d\tau = b^{(fem)}(\tau)d\tau \int_{a_1}^{a_2} S^{(fem)}(x)\phi_x e^{-\rho x}dx, \quad (6.61)$$

gdzie $[a_1, a_2)$ oznacza przedział wieku prokreacyjnego kobiet (przyjmujemy $a_1 = 15$, $a_2 = 50$ lat).

Otrzymaliśmy w ten sposób tzw. **integralne równanie reprodukcji ludności**, które można uprościć do postaci

$$1 = \int_{a_1}^{a_2} S^{(fem)}(x)\phi_x e^{-\rho x}dx, \quad (6.62)$$

zwanego równaniem charakterystycznym. Równanie to wskazuje na relację wiążącą umieralność, reprezentowaną przez funkcję przeżycia $S(x)$, funkcję płodności ϕ_x oraz istotną stopę ρ wzrostu populacji.

Wyrażenie $S^{(fem)}(x)\phi_x$ występujące pod całką w (6.62) nazywamy **funkcją płodności netto**. Będziemy ją oznaczać dalej symbolem Φ_x , czyli

$$\Phi_x = S^{(fem)}(x)\phi_x \quad \text{dla } x \in [a_1, a_2). \quad (6.63)$$

Podobne rozumowanie, jak to prowadzące do formuły (6.61), możemy zastosować w odniesieniu do liczby zgonów w wieku $[x, x + dx)$, czyli do wyrażenia $d_x(\tau)dx$, które przez analogię do (6.60) jest równe

$$S(x)b(\tau)\mu_x e^{-\rho x}dx, \quad (6.64)$$

gdzie μ_x oznacza intensywność zgonów w danej populacji dla ustalonego wieku x , niezależną od czasu τ .

Całka oznaczona z (6.64) dla $x \in [0, \infty)$ prowadzi do funkcji $d(\tau)$, stąd iloczyn $d(\tau)d\tau$, tj. łączna liczba zgonów w okresie $[\tau, \tau + d\tau)$, wynosi

$$d(\tau)d\tau = b(\tau)d\tau \int_0^{\infty} S(x)\mu_x e^{-\rho x} dx. \quad (6.65)$$

6.5.3. Oczekiwana długość generacji

Koncepcję „długości generacji” wyjaśnić można za pomocą hipotetycznej sytuacji, w której wszystkie urodzenia dzieci wydanych na świat przez kobiety z generacji s mają miejsce w tym samym, wspólnym momencie, podobnie urodzenia z generacji $s + 1$ itd. Wówczas odstęp czasu pomiędzy momentami urodzeń matek i ich dzieci nazwać możemy długością generacji. W rzeczywistości urodzenia dzieci są rozciągnięte na przestrzeni całego okresu rozrodczego kobiet, a więc określenie długości generacji w sposób punktowy nie jest możliwe. Możemy jednak rozważać tzw. **średnią długość generacji**.

Przez średnią długość generacji T (*mean length of a generation*) w linii żeńskiej rozumieć będziemy środek ciężkości funkcji płodności netto Φ_x określonej wzorem (6.63). Mamy więc

$$T = \frac{\int_{a_1}^{a_2} x \cdot \Phi_x dx}{\int_{a_1}^{a_2} \Phi_x dx} = \frac{\int_{a_1}^{a_2} x S^{(fem)}(x) \phi_x dx}{\int_{a_1}^{a_2} S^{(fem)}(x) \phi_x dx}, \quad (6.66)$$

gdzie $[a_1, a_2)$ określa ustalony przedział wieku rozrodczego, tutaj przyjmujemy $a_1 = 15, a_2 = 50$ lat.

W ogólnym przypadku tak zdefiniowany parametr T nie jest równy średniemu wiekowi matek w chwili pogoju. Istotnie, np. w modelu ludności ustabilizowanej średni wiek matek $\bar{x}_u^{(fem)}$ w momencie urodzenia ich dzieci jest na mocy (6.60) oraz (6.63) równy

$$\bar{x}_u^{(fem)} = \frac{\int_{a_1}^{a_2} x \cdot \Phi_x \cdot e^{-\rho x} dx}{\int_{a_1}^{a_2} \Phi_x \cdot e^{-\rho x} dx} = \frac{\int_{a_1}^{a_2} x S^{(fem)}(x) \phi_x e^{-\rho x} dx}{\int_{a_1}^{a_2} S^{(fem)}(x) \phi_x e^{-\rho x} dx}. \quad (6.67)$$

Formuła ta nie pokrywa się z (6.66) z wyjątkiem przypadku, gdy $\rho = 0$, czyli dla populacji stacjonarnej.

Metodę szacowania parametru T odnieść możemy do dyskretnego odpowiednika prawej strony (6.66), czyli

$$T \approx \frac{\sum_{x=a_1}^{a_2-1} (x + \frac{1}{2}) \cdot {}_{x+\frac{1}{2}}p_0 \cdot f_x}{\sum_{x=a_1}^{a_2-1} {}_{x+\frac{1}{2}}p_0 \cdot f_x}, \quad (6.68)$$

gdzie ${}_{x+\frac{1}{2}}p_0$ oznacza prawdopodobieństwo dożycia przez noworodka płci żeńskiej wieku $x + \frac{1}{2}$ (wyznaczone z tablic trwania życia dla kobiet), natomiast f_x jest cząstkowym współczynnikiem płodności dla grupy wieku $[x, x + 1)$.

W przypadku, gdy dysponujemy współczynnikami płodności ${}_n f_x$ w szerszych grupach wieku $[x, x + n)$, tj. dla $n > 1$, formułę szacowania parametru T możemy zapisać wzorem [Brown 1997, s. 153]

$$T \approx \frac{\sum_x (x + \frac{n}{2}) \cdot {}_{x+\frac{n}{2}}p_0 \cdot {}_n f_x}{\sum_x {}_{x+\frac{n}{2}}p_0 \cdot {}_n f_x}, \quad (6.69)$$

Zauważymy także, że jeśli T jest średnią długością generacji w linii żeńskiej, to iloraz urodzeń żeńskich $B^{(fem)}(t)$ oraz $B^{(fem)}(t - T)$ możemy przyjąć jako aproksymację współczynnika reprodukcji netto NRR , czyli

$$NRR \approx \frac{B^{(fem)}(t)}{B^{(fem)}(t - T)}. \quad (6.70)$$

6.5.4. Wyznaczanie współczynnika przyrostu naturalnego

Ważnym zagadnieniem w modelu Lotki jest problem rozwiązania równania charakterystycznego (6.62) ze względu na współczynnik przyrostu naturalnego ρ w danych warunkach płodności i umieralności zdefiniowanych funkcjami ϕ_x i μ_x .

Lotka udowodnił, że równanie (6.62) ma dokładnie jeden pierwiastek rzeczywisty. Pozostałe pierwiastki ρ_i , poza pierwszym ρ_1 , są liczbami zespolonymi, przy czym zachodzi nierówność $Re(\rho_i) < \rho_1$, gdzie $Re(z)$ oznacza część rzeczywistą liczby zespolonej z . Jednak jego rozwiązanie wymaga znajomości funkcji ϕ_x i μ_x , podczas gdy w praktyce dysponujemy zazwyczaj tylko ich aproksymacjami, tj. cząstkowymi współczynnikami płodności oraz współczynnikami zgonów. Istotną stopę wzrostu ρ można wtedy przybliżyć, rozwiązując np. wersję dyskretną równania (6.62) lub też na podstawie relacji łączącej ten współczynnik ze współczynnikiem reprodukcji netto NRR . Poniżej przedstawiona zostanie bliżej druga z wymienionych metod.

Niech T oznacza oczekiwaną długość generacji. Rozważmy model ludności ustabilizowanej z istotną stopą wzrostu ρ . Niech $b^{(fem)}(\tau)d\tau$ będzie liczbą urodzeń żeńskich w okresie $[\tau, \tau + d\tau)$, przy czym dla funkcji $b^{(fem)}(\tau)$ ma miejsce zależność (6.46), czyli

$$b^{(fem)}(\tau) = b^{(fem)}(\tau - T)e^{\rho T}. \quad (6.71)$$

Po scałkowaniu dwustronnie na przedziale $[\tau_0, \tau_0 + 1]$ otrzymujemy

$$\begin{aligned} B^{(fem)}(t) &= \int_{\tau_0}^{\tau_0+1} b^{(fem)}(\tau)d\tau = e^{\rho T} \int_{\tau_0}^{\tau_0+1} b^{(fem)}(\tau - T)d\tau = \\ &= e^{\rho T} \int_{\tau_0-T}^{\tau_0+1-T} b^{(fem)}(\tau)d\tau = e^{\rho T} B^{(fem)}(t - T), \end{aligned} \quad (6.72)$$

gdzie $t = [\tau_0, \tau_0 + 1]$, $t - T = [\tau_0 - T, \tau_0 + 1 - T]$, natomiast $B^{(fem)}(t)$ jest liczbą urodzeń żeńskich w roku $t = [\tau_0, \tau_0 + 1]$.

Z (6.72) oraz (6.70) mamy

$$NRR \approx \exp\{\rho T\}, \quad (6.73)$$

a następnie

$$\rho \approx \frac{\ln(NRR)}{T}. \quad (6.74)$$

Korzystając z faktu, że $e^\rho \approx 1 + \rho$, możemy zastosować dalszą aproksymację w postaci

$$e^{\rho T} \approx (1 + \rho)^T, \quad (6.75)$$

co prowadzi do przybliżonej formuły na rzeczywisty współczynnik przyrostu naturalnego

$$\rho \approx \sqrt[T]{NRR} - 1. \quad (6.76)$$

Inną, iteracyjną metodę aproksymacji istotnej stopy wzrostu ρ znaleźć można m.in. w [Keyfitz 1968].

W celu oszacowania współczynnika ρ na podstawie formuły (6.74) lub (6.76) szacujemy najpierw współczynnik reprodukcji netto NRR oraz średnią długość generacji T .

Przykład 22. Na podstawie danych z przykładu 16 obliczymy rzeczywisty współczynnik przyrostu naturalnego. W tym celu wyznaczmy najpierw średnią długość generacji T na podstawie formuły (6.69). Obliczenia pośrednie zawiera tablica 17.

Tablica 17. Prawdopodobieństwa ${}_{x+2,5}p_0$, współczynniki płodności ${}_5f_x$ oraz obliczenia pośrednie parametru T

x	${}_{x+2,5}p_0$	${}_5f_x$	${}_{x+2,5}p_0 \cdot {}_5f_x$	$(x+2,5) \cdot {}_{x+2,5}p_0 \cdot {}_5f_x$
(1)	(2)	(3)	(4)	(5)
15	0,9944	0,0109	0,0108	0,1892
20	0,9933	0,0506	0,0503	1,1312
25	0,9920	0,0991	0,0983	2,7037
30	0,9904	0,0861	0,0853	2,7724
35	0,9879	0,0372	0,0367	1,3780
40	0,9838	0,0075	0,0074	0,3130
45	0,9767	0,0003	0,0002	0,0154
Razem			0,2891	8,5030

Źródło: obliczenia na podstawie danych GUS

Iloraz sumy wyrazów w kolumnie 5 przez sumę wyrazów w kolumnie 4 daje oszacowanie parametru $T = 8,5030/0,2891 = 29,41$. Ponieważ współczynnik reprodukcji netto wynosi w tym przypadku 0,701 [por. przykład 16, s. 120 – 121], zatem na podstawie przybliżenia (6.74) otrzymujemy $\rho \approx -0,0121$, a na podstawie przybliżenia (6.76) mamy $\rho \approx -0,0120$.

Stopa przyrostu naturalnego w roku 2017 w Polsce wynosiła ok. $-0,00002$, co oznacza ubytek ludności rzędu ok. 2 osób na 100 000 mieszkańców [por. przykład 18, s. 125 – 126]. Wartość ta różni się od istotnej stopy wzrostu $-0,0120$. Druga z tym wartości ma charakter hipotetyczny. Informuje, ile wynosiłaby stopa przyrostu naturalnego, gdyby intensywność płodności i umieralności obserwowane w roku 2017 utrzymywały się niezmiennie na tym samym poziomie przez długi okres czasu.

6.5.5. Intensywność wzrostu a intensywność urodzeń, zgonów

Określmy intensywność urodzeń $\beta(\tau)$ oraz zgonów $\mu(\tau)$ w chwili τ . W pierwszym przypadku mamy

$$\beta(\tau) = \frac{B(\tau)}{B(\tau) \int_0^\infty S(x)e^{-\rho x} dx} = \frac{1}{\int_0^\infty S(x)e^{-\rho x} dx}, \quad (6.77)$$

gdzie $B(\tau)d\tau$ oznacza liczbę urodzeń w przedziale $[\tau, \tau+d\tau)$, natomiast wielkość populacji $L(\tau)$ określona jest przez formułę analogiczną do (6.52).

Podobnie, intensywność zgonów w chwili τ jest równa

$$\mu(\tau) = \frac{D(\tau)}{B(\tau) \int_0^\infty S(x)e^{-\rho x} dx} = \frac{\int_0^\infty S(x)e^{-\rho x} \mu_x dx}{\int_0^\infty S(x)e^{-\rho x} dx}. \quad (6.78)$$

Zauważymy, że prawe strony (6.77) oraz (6.78) nie zależą od czasu τ , czyli obie wielkości są stałe w czasie. Oznaczmy je dalej symbolami β oraz μ i nazwiemy odpowiednio **istotną stopą urodzeń** i **istotną stopą zgonów**. Mamy zatem

$$\beta = \frac{1}{\int_0^\infty S(x)e^{-\rho x} dx}, \quad (6.79)$$

$$\mu = \beta \int_0^\infty S(x)e^{-\rho x} \mu_x dx. \quad (6.80)$$

Po scałkowaniu przez części prawej strony (6.80) mamy

$$\begin{aligned} \mu &= \beta \int_0^\infty S(x)e^{-\rho x} \mu_x dx = \\ &= -\beta e^{-\rho x} S(x) \Big|_0^\infty - \rho \cdot \beta \int_0^\infty e^{-\rho x} S(x) dx = \\ &= \beta - \rho \cdot \beta \int_0^\infty e^{-\rho x} S(x) dx. \end{aligned} \quad (6.81)$$

Wyrażenie $\beta \int_0^\infty e^{-\rho x} S(x) dx$ równe jest 1, co wynika z (6.62). Wobec tego mamy dalej

$$\mu = \beta - \rho \cdot 1 = \beta - \rho, \quad (6.82)$$

co prowadzi ostatecznie do równości

$$\rho = \beta - \mu, \quad (6.83)$$

a zatem istotna stopa wzrostu ρ w modelu ludności ustabilizowanej jest różnicą istotnych stóp urodzeń i zgonów.

6.5.6. Średnie trwanie życia a średni wiek w chwili zgonu

Pokażemy na przykładzie modelu Lotki, że średnie trwanie życia noworodka \dot{e}_0 i średni wiek \bar{x} w chwili zgonu nie są tożsame i nie należy mylić tych pojęć.

Średni wiek \bar{x} w chwili zgonu w modelu ludności ustabilizowanej wyznaczmy jako całkę z iloczynu wieku x i liczby zgonów $D_x(\tau)dx d\tau$ w grupie wieku $[x, x + dx]$ i w okresie $[\tau, \tau + d\tau]$, podzieloną przez łączną liczbę zgonów $D(\tau)d\tau$ w tym okresie, co prowadzi do formuły

$$\begin{aligned}\bar{x} &= \frac{\int_0^\infty x \cdot D_x(\tau)dx}{D(\tau)} = \frac{B(\tau) \int_0^\infty x \cdot S(x)e^{-\rho x} \mu_x dx}{B(\tau) \int_0^\infty S(x)e^{-\rho x} \mu_x dx} = \\ &= \frac{\int_0^\infty x \cdot S(x)e^{-\rho x} \mu_x dx}{\int_0^\infty S(x)e^{-\rho x} \mu_x dx}.\end{aligned}\tag{6.84}$$

Wynik ten różni się od średniego trwania życia \dot{e}_0 , które jest wartością oczekiwaną czasu życia noworodka. Mamy bowiem na podstawie (4.23)

$$\dot{e}_0 \equiv E(X) = \int_0^\infty xS(x)\mu_x dx,\tag{6.85}$$

gdzie $f_X(x)$ oraz μ_x oznaczają odpowiednio funkcję gęstości rozkładu czasu życia X oraz funkcję intensywności zgonów.

Dodatkowo, z własności funkcji gęstości zachodzi równość

$$\int_0^\infty f_X(x)dx = \int_0^\infty S(x)\mu_x dx = 1.\tag{6.86}$$

Tym samym (6.85) możemy zapisać równoważnie jako

$$\dot{e}_0 = \frac{\int_0^\infty xS(x)\mu_x dx}{\int_0^\infty S(x)\mu_x dx}.\tag{6.87}$$

Zauważmy, że w ogólnym przypadku prawa strona (6.84) nie jest równa prawej stronie (6.87), czyli

$$\bar{x} \neq \dot{e}_0.\tag{6.88}$$

Wyjątek stanowi model ludności stacjonarnej, gdy $\rho = 0$.

Ćwiczenia

Ćwiczenie 6.1

Obliczyć, po jakim czasie populacja zwiększy się dwukrotnie, jeśli charakteryzuje ją stała stopa wzrostu równa:

1. $r = 0,01$,
2. $r = 0,05$.

Ćwiczenie 6.2

Na podstawie danych GUS o liczbie ludności Polski na koniec lat 2003, 2010 i 2017, obliczyć średnie tempo wzrostu w okresie:

1. 2004 – 2010,
2. 2011 – 2017.

Ćwiczenie 6.3

W roku 2017 pewna populacja liczyła 36 mln ludności. Zakładając, że średnia stopa wzrostu tej populacji jest ujemna i wynosi $-0,001$ oraz utrzyma się na tym poziomie w przyszłych okresach, obliczyć:

1. Liczbę ludności w roku 2030,
2. Liczbę ludności w roku 2050,
3. Za ile lat ludność tej populacji zmniejszy się o $1/4$?

Ćwiczenie 6.4

Na podstawie wyników obliczeń współczynnika NRR dla roku 2000 opartych na danych z tablicy 16 (strona 121) oraz liczb dożywających l_x zaczerpniętych z tablicy trwania życia dla kobiet z roku 2000 (stat.gov.pl), obliczyć średnią długość generacji T w linii żeńskiej oraz istotną stopę wzrostu populacji ρ . Wyniki porównać z rezultatami otrzymanymi w przykładzie 22 (s. 143 – 144).

Bibliografia

- Bolesławski L., (1973), *Budowa tablic trwania życia. Teoria i praktyka*, GUS, Warszawa.
- Brown R., (1997), *Introduction to the Mathematics of Demography*, Third Edition, ACTEX Publications Inc.
- Caselli G., Vallin J., Wunsch G., (2006), *Demography: Analysis and Synthesis. Four Volume Set: A Treatise in Population*, Elsevier, London–New York–Tokyo.
- Frątczak E., (1997), *Analiza historii zdarzeń – elementy teorii, wybrane przykłady zastosowań z wykorzystaniem pakietu TDA*, SGH, Warszawa.
- Holzer Z., (2006), *Demografia*, PWE, Warszawa.
- Lotka A. J., (1939), *A Contribution to the Theory of Self-Renewing Aggregates*, The Annals of Mathematical Statistics, vol. 10, 1–25.
- Keyfitz N., (1968), *Introduction to the Mathematics of Population*, Addison–Wesley, Reading Mass.
- Kędelski M., Paradysz J., (2006), *Demografia*, Wyd. AE, Poznań.
- Kurkiewicz J. (red.), (2010), *Procesy demograficzne i metody ich analizy*, Wyd. UE w Krakowie.
- Okólski M., (2004), *Demografia. Podstawowe pojęcia i teorie w encyklopedycznym zarysie*, Wyd. Scholar, Warszawa.
- Pressat R. (1966), *Analiza demograficzna – metody wyniki, zastosowania*, PWN, Warszawa.
- Preston S. H., Heuveline P., Guillot M., (2001), *Demography. Measuring and Modeling Population Processes*, Blackwell Publishing Ltd., Malden–Oxford–Carlton.
- Skałba M., (2002), *Ubezpieczenia na życie*, WNT, Warszawa.
- Tabeau E., Berg Jehs A., Heathcote Ch. (ed.), (2001), *Forecasting Mortality in Developed Countries. Insights from Statistical, Demographic and Epidemiological Perspective*, Kluwer Academic Publishers, London.
- Tablice trwania życia*, (2017), Wyd. GUS.

