

DETEKCJA ZMIANY DRYFU W MODELOWANIU NATEŻENIA ŚMIERTELNOŚCI¹

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY
Nr 15(21)

Michał Krawiec

Uniwersytet Wrocławski

Zbigniew Palmowski

Politechnika Wroclawska

e-mails: michal.krzysztof.krawiec@gmail.com;

zbigniew.palmowski@gmail.com

ISSN 1644-6739
e-ISSN 2449-9765

DOI: 10.15611/sps.2017.15.07

JEL Classification: C01, C02, C60, C80, G22

Streszczenie: Wydłużający się czas trwania życia stawia nowe wyzwania w różnych obszarach ekonomicznych i ma kluczowy związek z obowiązującym systemem emerytalnym. W celu ograniczenia ryzyka związanego ze starzejącym się społeczeństwem potrzebne jest określenie, jak bardzo ten czas się wydłuża. Celem artykułu jest badanie zmian zachodzących w trendzie procesu natężenia śmiertelności poprzez zagadnienie optymalnej detekcji. Na podstawie konstrukcji tzw. uogólnionej statystyki Shiryayeva-Robertsza zbadano, jak zmieniano się natężenie śmiertelności populacji Polski w latach 1990-2014. W artykule logarytm z natężenia śmiertelności modelowany jest ruchem Browna, któremu w pewnym losowym, nieobserwowalnym momencie dochodzi dryf. Skonstruowano optymalny moment zatrzymania i opisano algorytm detekcji zmiany dryfu w wersji dyskretnej. Przedstawiono kalibrację modelu oraz przeanalizowano dane dostarczane przez GUS. Narysowano też wykresy i sporządzono wnioski dotyczące modelu detekcji oraz jego parametrów.

Słowa kluczowe: natężenie śmiertelności, detekcja, statystyka Shiryayeva-Robertsza, ruch Browna, tablice trwania życia.

1. Wstęp

Na świecie stale obserwuje się wydłużanie się średniego czasu życia ludzkości. Jest to niewątpliwie związane z rozwojem cywilizacji, medycyny, z coraz większą wiedzą o człowieku. Podobna sytuacja ma miejsce w Polsce – ludzie w naszym kraju również żyją coraz dłużej. Stawia to wiele wyzwań przed dynamicznie rozwijającym się rynkiem ubezpieczeń i systemów emerytalnych. Produkty ubezpieczeniowe dotyczą przyszłości, której nikt nie jest w stanie przewidzieć, trzeba ją

¹ Praca częściowo finansowana z grantu NCN nr 2015/17/B/ST1/01102.

jednak w pewien sposób modelować i prognozować. Stosowane modele nie mogą opierać się jedynie na dzisiejszej wiedzy – muszą być na bieżąco dostosowywane do zmieniającego się w trakcie lat świata. Jednym z podstawowych powodów, dla których modele muszą być dostosowywane, jest zmieniające się co pewną liczbę lat tempo wydłużania średniego czasu życia.

Na tym właśnie tempie skupiono się w niniejszym artykule. Przyjrano się, jak zmieniało się natężenie śmiertelności populacji Polski w ciągu ostatnich 25 lat, i przedstawiono metodę pozwalającą wykrywać okresy coraz szybszego zmniejszania śmiertelności oraz okresy stabilizacji. Określenie, jak bardzo wydłuża się czas życia, a tym samym zmniejsza natężenie śmiertelności, jest kluczowym wyzwaniem zarówno polskiej, jak i światowej gospodarki. Zrozumienie tego procesu jest niezmiernie istotne między innymi w celu poprawnej konstrukcji instrumentów finansowo-aktuarialnych, jak również jest ściśle związane z systemami emerytalnymi. Aby uzmysłowić sobie powagę problemu, wystarczy powołać się na przypadek Francji, która w 2006 roku dokonała korekty krajowych tablic trwania życia, stworzonych w roku 1998, co spowodowało potrzebę wzrostu o 8% planowanych wydatków w systemie emerytalnym.

W artykule zaprezentowano algorytm detekcji, oparty na uogólnionej wersji tzw. statystyki Shiryaeva-Roberts'a, który będzie użyty w celu wykrycia zmian tendencji (dryfu) w modelowanym procesie natężenia śmiertelności. Będzie się obserwować, jak dla ustalonego wieku x lat zmieniała się jego intensywność śmiertelności podczas lat 1990-2014 na podstawie tablic trwania życia udostępnianych przez Główny Urząd Statystyczny i na podstawie tych danych pokazane zostanie, jak działa algorytm detekcji. Uwaga zostanie również zwrócona na istotność pewnych parametrów wchodzących w skład omawianego modelu.

Samo zagadnienie detekcji sięga lat 50. – wówczas to Kołmogorow wraz z innymi rosyjskimi naukowcami rozważał jego zastosowanie w problemach inżynierskich, m.in. do oddzielania nadchodzącego sygnału od szumu. Sformułowano wtedy wersję tego zagadnienia opartą na ruchu Browna, któremu w pewnej losowej chwili dochodzi dryf. Artykuł ten oparto na podobnym sformułowaniu problemu od strony matematycznej, jednak innym – aktuarialnym – zastosowaniu. Optymalne rozwiązanie problemu (wraz z pełnym, matematycznym sformułowaniem i wyprowadzeniem) w swoich pracach podaje Shiryaev [1978; 2006].

W dzisiejszych czasach zagadnienie optymalnej detekcji jest coraz szerzej uogólniane od strony matematycznej. Badane są jego różne sformułowania; często zamiast ruchu Browna stosowany jest proces Poissona, również złożony. Analizowane są także różne kryteria optymalności rozwiązania (oprócz podejścia Bayesowskiego, jak w tym artykule, czasami rozważa się też podejście minimaksowe). Przykładowe prace na ten temat to [Peskir, Shiryaev 2002; Gapeev 2005; El Karoui, Loisel 2015].

2. Sformułowanie problemu

Przedmiotem badań jest natężenie śmiertelności μ_t , o którym zakłada się, że jest pewnym obserwowanym przez nas procesem losowym. W praktyce model natężenia μ_t może w ciągu lat ulegać zmianom. Moment wystąpienia istotnej zmiany „charakteru” tego procesu będzie oznaczany przez θ . Chwila ta zazwyczaj będzie losowa. W skrócie będzie ona nazywana momentem zmiany procesu. Można więc zapisać natężenie śmiertelności w następującej postaci:

$$\mu_t = \begin{cases} \mu_t^{(1)} & t < \theta, \\ \mu_t^{(2)} & t \geq \theta, \end{cases}$$

gdzie $t \geq 0$. Przyjęte zostanie, że natężenie to składa się z nielosowej, deterministycznej części $\bar{\mu}_t$ – niezmiennej, tzn. tej samej dla $\mu_t^{(1)}$ i $\mu_t^{(2)}$, oraz losowej perturbacji, której „charakter” ulega zmianie w momencie zmiany θ . Aby móc uzyskać wartości ujemne w modelowanym przez nas procesie perturbacji, na natężenie śmiertelności nałożony zostanie logarytm. Zapisując symbolicznie:

$$\log \mu_t^{(1)} = \log \bar{\mu}_t + X_t^{(1)}, \tag{1}$$

$$\log \mu_t^{(2)} = \log \bar{\mu}_t + X_t^{(2)},$$

gdzie $X_t^{(1)}$ i $X_t^{(2)}$ są losowymi perturbacjami, będzie się zakładać, że przed momentem zmiany średnia wartość perturbacji wynosi 0, a po zmianie będzie ona równa $r \neq 0$. W języku wartości oczekiwanych można zapisać: $EX_t^{(1)} = 0$ oraz $EX_t^{(2)} = rt$. Podkreślić należy, że przez proces $X = (X_t)_{t \geq 0}$ jest rozumiany łączny (ulegający zmianie w chwili θ) proces perturbacji, tj.:

$$X_t = \begin{cases} X_t^{(1)} & t < \theta, \\ X_t^{(2)} & t \geq \theta. \end{cases} \quad (2)$$

O przedstawionym problemie należy myśleć w ten sposób: obserwuje się natężenie śmiertelności dla ustalonego wieku x lat (np. dla pięćdziesięciolatków) zmieniające się poprzez lata (czyli w roku 1990, 1991, 1992, ...) i chce się stworzyć metodę statystyczną, która pozwoli w optymalny sposób wybrać moment, w którym „zachowanie” tego obserwowanego procesu ulega zmianie. Powstaje zatem pytanie: jak w tym kontekście rozumieć „optymalny sposób”?

Jedno z podstawowych kryteriów zaproponowanych przez Shiryaeva uwzględnia zarówno średnie opóźnienie wykrycia zmiany, jak i prawdopodobieństwo „fałszywego alarmu”. Będzie więc rozważne kryterium następującej postaci:

$$B(c) = \inf_{\tau} \{P(\tau < \theta) + cE(\tau - \theta)^+\} \quad (3)$$

i poszukiwane takie τ^* , które będzie czasem zatrzymania realizującym powyższe infimum. Ten moment będzie nazywany optymalnym czasem zatrzymania (innymi słowy τ^* opisuje strategię, na podstawie której podczas obserwacji procesu X można zdecydować o „zgłoszeniu alarmu”, a zatem o tym, że „charakter” obserwowanego procesu uległ zmianie).

W tej pracy procesy $X_t^{(1)}$ oraz $X_t^{(2)}$ są procesami Lévy’ego, tj. procesami o stacjonarnych i niezależnych przyrostach. Klasycznym przykładem takiego procesu jest ruch Browna i ten przykład zostanie użyty do analizy numerycznej w dalszej części artykułu, tzn.:

$$X_t = \begin{cases} \sigma W_t & t < \theta, \\ \sigma W_t + r(t - \theta) & t \geq \theta, \end{cases} \quad (4)$$

gdzie W_t jest standardowym ruchem Browna, σ jest parametrem zmienności (odchylenia standardowego), a r stałą odpowiadającą za dryf. Rozważany będzie czas $t \geq 0$.

Od tego momentu autorzy będą się skupiać na tym właśnie przykładzie problemu detekcji, oznaczonym numerem (4).

Wprowadzone zostaną podstawowe założenia dotyczące momentu wystąpienia zmiany procesu θ . Zastosowane zostanie podejście Bayesowskie, w którym jest to zmienna losowa o zadanym rozkładzie *a priori*. Po pierwsze, zostanie dopuszczona możliwość wystąpienia momentu zmiany już od początku z pewnym prawdopodobieństwem:

$$P(\theta = 0) = \pi, \quad \pi \in [0,1].$$

Po drugie, jeżeli moment zmiany nie wystąpi w chwili 0, autorzy chcą, aby miał on rozkład wykładniczy z parametrem λ , tzn.:

$$P(\theta > t \mid \theta > 0) = e^{-\lambda t}.$$

Kluczową rolę w rozważaniach odegra prawdopodobieństwo *a posteriori* (względem obserwacji procesu X_t), które oznaczono jako:

$$\pi_t = P(\theta \leq t \mid F_t^X),$$

gdzie przez F_t^X rozumiana jest historia procesu X_t od chwili 0 do chwili t .

Można pokazać, że używając powyższego prawdopodobieństwa *a posteriori* π_t , kryterium optymalności sformułowane w (3) da się przeformułować. Przyjmuje wtedy następującą formę:

$$B_\pi(c) = \inf_\tau E_\pi \left[(1 - \pi_\tau) + c \int_0^\tau \pi_s ds \right], \quad (5)$$

gdzie π w indeksie dolnym oznacza, że $\pi_0 = P(\theta = 0) = \pi$. Tej formy będziemy używać w dalszej części tekstu.

Wprowadzone zostanie jeszcze kilka dodatkowych oznaczeń. Oprócz miary probabilistycznej P , używanej powyżej, na tej samej przestrzeni (oznaczono ją standardowo (Ω, \mathcal{F})) będzie się rozróżniać też takie miary (probabilistyczne), w których moment wystąpienia zmiany θ jest ustalony i nielosowy. Miary te oznaczone będą przez P_s , gdzie s może być dowolną nieujemną liczbą rzeczywistą lub nawet nieskończonością – wtedy $P_s(\theta = s) = 1$. Przede wszystkim wyróżnia się dwie z tych miar: P_0 i P_∞ . Pierwsza z nich oznacza, że zaburzenie wystąpiło już w chwili 0, a więc już od początku obserwacji procesu X_t widać tak naprawdę proces $X_t^{(2)}$. Analogicznie zgodnie z miarą P_∞ zaburzenie nigdy nie następuje (tzn. nastąpi dopiero „w nieskończoności”), a więc przez cały czas obserwacji mamy do czynienia z procesem $X_t^{(1)}$.

W problemie detekcji niezwykle istotną rolę odgrywa iloraz wiarygodności tych dwóch szczególnych miar. Będzie on oznaczany przez L_t :

$$L_t = \frac{dP_0}{dP_\infty} \Big|_{\mathbb{F}_t^X},$$

gdzie \mathbb{F}_t^X jest naturalną filtracją procesu X . Na tym ilorazie wiarogodności oparty jest jeszcze jeden bardzo istotny proces, który jest powiązany z tzw. statystyką Shiryaeva-Robertsza, o której mowa w dalszej części tekstu. Proces ten będzie oznaczony przez ϕ_t i definiowany jako:

$$\phi_t = \frac{\pi_t}{1 - \pi_t}. \quad (6)$$

Można pokazać, że powyższy proces można zapisać w następującej postaci:

$$\phi_t = e^{\lambda t} L_t \left(\phi_0 + \lambda \int_0^t \frac{e^{-\lambda s}}{L_s} ds \right), \quad \phi_0 = \frac{\pi}{1 - \pi}. \quad (7)$$

W tym miejscu warto napisać, ile wynoszą L_t i ϕ_t w rozważanym modelu (4). Mamy mianowicie:

$$L_t = e^{H_t}, \quad H_t = \frac{r}{\sigma^2} X_t - \frac{r^2}{2\sigma^2} t,$$

$$\phi_t = \phi_0 \cdot e^{\lambda t + H_t} + \lambda \int_0^t e^{-(\lambda s + H_s)} ds.$$

Zauważyć należy jeszcze w tym miejscu, że dla wszystkich $t \geq 0$ mamy $\pi_t \in [0, 1]$, a wtedy (wyłączając przypadek $\pi_t = 1$) wzór (6) można odwrócić, otrzymując:

$$\pi_t = \frac{\phi_t}{1 + \phi_t}. \quad (8)$$

Obserwacja ta będzie przydatna w dalszych rozważaniach.

3. Uogólniona Statystyka Shiryaeva-Robertsza

Głównym celem tej pracy jest konstrukcja i zastosowanie uogólnionej statystyki Shiryaeva-Robertsza (w skrócie będziemy ją oznaczać przez USR). Zacząć należy od tego, czym jest klasyczna statystyka SR. W literaturze statystycznej jest ona zapisywana jako:

$$\psi_t = \int_0^t \frac{L_t}{L_s} ds.$$

Zauważyć trzeba, że można ją otrzymać z wprowadzonej w (7) formy ϕ_t . Mianowicie przy założeniu, że $\pi = \pi_0 = 0$, otrzymuje się też $\phi_0 = 0$, a wtedy:

$$\lim_{\lambda \rightarrow 0} \frac{\phi_t}{\lambda} = \int_0^t \frac{L_t}{L_s} ds = \psi_t.$$

W tej pracy autorzy chcą przyjąć ogólniejszą postać tej statystyki i konstrukcję USR oparli właśnie na ϕ_t . Należy przyjrzeć się bliżej ilorazowi L_t . Trzeba ustalić chwilę t i podzielić odcinek $[0, t]$ na n równych części punktami $0 = t_0 < t_1 < t_2 < \dots < t_n = t$. Można wtedy zapisać:

$$L_{t_n} = e^{\frac{r}{\sigma^2} \cdot X_0} \cdot \prod_{i=1}^n e^{\frac{r}{\sigma^2} \cdot (X_{t_i} - X_{t_{i-1}}) - \frac{r^2}{2\sigma^2} \cdot (t_i - t_{i-1})}.$$

Dla prostoty zapisu przyjęto, że dla wszystkich $i \in \{1, \dots, n\}$ $t_i - t_{i-1} = 1$, tzn. $t_0 = 0, t_1 = 1, t_2 = 2, \dots, t_n = n$. Przyjęto też oznaczenie dla przyrostów $x_i = X_i - X_{i-1}$ dla wszystkich $i \in \{1, \dots, n\}$ oraz że $X_0 = 0$. Wtedy można zapisać krócej:

$$L_n = \prod_{i=1}^n e^{\frac{r}{\sigma^2} x_i - \frac{r^2}{2\sigma^2}}.$$

Dla $i = 0$ przyjęto, że $L_0 = 1$. Można zauważyć, że L_n w takiej postaci faktycznie jest ilorazem wiarygodności – łatwo da się sprawdzić, że oznaczając przez $f(x)$ i $g(x)$ gęstości rozkładów normalnych odpowiednio $N(r, \sigma^2)$ i $N(0, \sigma^2)$, otrzyma się dokładnie:

$$L_n = \prod_{i=1}^n \frac{f(x_i)}{g(x_i)}.$$

Rozkłady te odpowiadają rozkładowi przyrostów procesu X_t w miarach odpowiednio P_0 i P_∞ (w jednostce czasu). Otrzymano zatem

L_n dla $n \in N$, które jest dyskretną wersją L_t dla $t \geq 0$. Podobnie można wyprowadzić dyskretną wersję dla ϕ_t :

$$\begin{aligned}\phi_n &= e^{\lambda n} L_n \left(\phi_0 + \lambda \sum_{k=0}^{n-1} \frac{e^{-\lambda k}}{L_k} \right) = \\ &= \phi_0 \cdot e^{\lambda n} \prod_{i=1}^n \frac{f(x_i)}{g(x_i)} + \lambda \sum_{k=0}^{n-1} e^{\lambda(n-k)} \prod_{i=k+1}^n \frac{f(x_i)}{g(x_i)}.\end{aligned}\quad (9)$$

Powyższą statystykę można zapisać również w postaci rekurencyjnej dla $n \geq 1$ (przy zadanym ϕ_0):

$$\phi_n = (\lambda + \phi_{n-1}) e^{\lambda} \frac{f(x_n)}{g(x_n)}.\quad (10)$$

Ta właśnie dyskretna wersja statystyki ϕ_n , $n \in N$, będzie nazywana uogólnioną statystyką Shiryaeva-Robertsza (USR) i będzie używana do analizy numerycznej w dalszej części artykułu.

Powracając do szukania optymalnego czasu zatrzymania dla problemu (3) (równoważnie (5)) w modelu (4), należy stwierdzić, że zgodnie ze standardową teorią znajdowania rozwiązań zagadnień optymalizacyjnych (przedstawioną szeroko w monografii [Peskir, Shiryaev 2006]) rozwiązanie tego zagadnienia sprowadza się do znalezienia funkcji $V(\pi)$ i punktu $A^* \in (0,1)$ spełniających układ:

$$\begin{aligned}AV(\pi) &= -c \cdot \pi, & \pi &\in [0, A^*), \\ V(\pi) &= 1 - \pi, & \pi &\in (A^*, 1],\end{aligned}\quad (11)$$

gdzie A oznacza generator procesu π_t . Do podania rozwiązania potrzebne jest zadanie pewnego dodatkowego warunku – bez wnikania w szczegóły (również są one częścią standardowej teorii) dla problemu sformułowanego w (4) jest to tzw. *smooth fit*:

$$V'(A^*) = -1.\quad (12)$$

Twierdzenie 1. *Optymalnym czasem zatrzymania dla problemu (3) w modelu (4) jest:*

$$\tau^* = \inf \{t \geq 0 : \pi_t \geq A^*\},$$

gdzie punkt $A^* \in (0,1)$ oraz optymalną funkcję wartości $V(\pi)$ należy wyznaczyć z układu (11) przy dodatkowym warunku (12).

Uwaga 1. *Twierdzenie to pozostaje prawdziwe również dla dużo ogólniejszych modeli postaci (2), gdzie \mathbf{A} jest tzw. pełnym generatorem, a warunek (12) może zostać zastąpiony tzw. continuous fit.*

Twierdzenie to przybliży już znacznie do przeprowadzenia całej procedury detekcji obserwowanego procesu. Mając bowiem wyprowadzoną statystykę USR (10) dla kolejnych chwil $n = 0, 1, 2, \dots$ korzystając z (8), można wyliczyć wartości π_n . Zgodnie z twierdzeniem optymalna detekcja („zgłoszenie alarmu”) następuje, gdy $\pi_n \geq A^*$, trzeba zatem jeszcze tylko znać próg A^* . Okazuje się, że wyznaczenie tego progu (rozwiązanie układu (11)) stanowi w ogólności trudne zagadnienie. Jednakże dla przyjętego w tym artykule modelu (4) zostało ono rozwiązane i podane przez Shiryaeva [1978; 2006]. Zgodnie z nim optymalna bariera A^* jest wyznaczana jako rozwiązanie równania:

$$C \int_0^{A^*} e^{-\Lambda(G(A^*)-G(u))} \frac{du}{u(1-u)^2} = 1,$$

gdzie:

$$G(u) = \ln \frac{u}{1-u} - \frac{1}{u}, \quad \Lambda = \frac{\lambda}{r^2/(2\sigma^2)}, \quad C = \frac{c}{r^2/(2\sigma^2)}.$$

Ponadto optymalną funkcją wartości jest:

$$V(\pi) = \begin{cases} (1 - A^*) - \int_{\pi}^{A^*} y(x) dx, & \pi \in [0, A^*], \\ 1 - \pi, & \pi \in [A^*, 1], \end{cases}$$

gdzie:

$$y(x) = -C \int_0^x e^{\Lambda(G(x)-G(u))} \frac{du}{u(1-u)^2}.$$

Otrzymano już całą procedurę, którą można stosować dla kolejnych chwil $n = 0, 1, 2, 3, \dots$ wiedząc, że znajduje ona optymalny czas zatrzymania w sensie rozważanego kryterium (tj. równoważnie (3) i (5)). Jedyne, co pozostało do zrobienia, aby móc stosować tę procedurę do analizy faktycznych danych, to kalibracja modelu.

4. Kalibracja modelu

Kalibracja modelu zależy przede wszystkim od charakteru procesu, który chce się analizować. Polega na dopasowaniu parametrów teoretycznego procesu do analizowanych danych. Inaczej będzie wyglądała kalibracja procesu ciągłego, skokowego oraz mieszanki obu. Dlatego też zamiast opisywać ogólne zasady kalibracji, skupiono się na rozpatrywanym w tym artykule modelu (4), czyli kalibracji ruchu Browna.

Na początek krótko o tym, jak w ogóle liczyć natężenie śmiertelności. Jak wiadomo, jest to wielkość określająca „tempo” wymierania określonej populacji w konkretnym wieku. Będzie ono liczone na podstawie tablic trwania życia (z których konkretnie – o tym w kolejnym rozdziale). Dla ustalonego wieku x lat, mając dane prawdopodobieństwo przeżycia roku p_x , natężenie śmiertelności będzie liczone, zakładając hipotezę przedziałami stałego natężenia zgonów (HCFM) jako $\mu_n = -\log p_x$, gdzie n odpowiada rokowi, dla którego sporządzono daną tablicę trwania życia. Należy zwrócić tu uwagę na nieco niestandardowe oznaczenie: dla modelowanego natężenia śmiertelności μ_n ustalony jest wiek x , a zmienną jest rok obserwacji n , natomiast prawdopodobieństwo p_x oznaczane jest standardowo i należy rozumieć przez to prawdopodobieństwo, że w danym ustalonym n -tym roku x -latek przeżyje rok; liczone jest jako l_{x+1} / l_x .

Niech będą dane obserwacje procesu μ_t do kalibracji – oznaczono je $\hat{\mu}_0, \hat{\mu}_1, \dots, \hat{\mu}_k$. Należy przez to rozumieć „dane historyczne” na kolejne lata, wartości procesu przed rozpoczęciem obserwacji w celu detekcji punktu zmiany (nieformalnie: przed chwilą $t = 0$). Odwołując się do postaci (1), trzeba przypomnieć dwie rzeczy. Po pierwsze, tak naprawdę modelowany ma być logarytm z natężenia śmiertelności, czyli używane będą wartości $\log \hat{\mu}_0, \dots, \log \hat{\mu}_k$. Po drugie, we wzorze występuje nielosowa składowa $\log \bar{\mu}_t$. W niej chciano zawrzeć wartość początkową $\log \mu_0$ (czyli wartość logarytmu natężenia śmiertelności na początku obserwacji w celu detekcji, czyli w chwili $t = 0$ – tak, aby spełnione było założenie $X_0 = 0$), oraz dryf – celem bowiem było, aby $EX_t^{(1)} = 0, t \geq 0$, a więc żeby „na początku” obserwowany proces X_t miał zerowy współczynnik dryfu. Żeby skalibrować dryf, policzono przyrosty powyższych logarytmów „historycznego” procesu:

$$m_i = \log \hat{\mu}_i - \log \hat{\mu}_{i-1}, \quad i = 1, \dots, k,$$

a następnie wyciągnięto z nich średnią:

$$m = \frac{1}{k} \sum_{i=1}^k m_i.$$

Tym samym otrzymuje się skalibrowany proces $\log \bar{\mu}_i = \log \mu_0 + m \cdot t$, który możemy odjąć od danych historycznych, otrzymując przetransformowane dane postaci:

$$\log \tilde{\mu}_i = \log \hat{\mu}_i - \log \mu_0 - m \cdot i,$$

których możemy już bezpośrednio użyć w celu kalibracji samego procesu perturbacji X_t .

Mówiąc o kalibracji ruchu Browna, należy podkreślić, że podstawowym parametrem do wyznaczenia jest jego odchylenie standardowe σ . Należy zauważyć, że w przyjętym modelu przyrosty procesu: $X_1 - X_0, X_2 - X_1, \dots$ mają rozkład normalny z wariancją równą σ^2 . Zatem estymator parametru σ zostanie wyznaczony na podstawie przyrostów obserwacji „historycznych” $\log \tilde{\mu}_i$, które oznaczone zostaną przez $y_i = \log \tilde{\mu}_i - \log \tilde{\mu}_{i-1}, i = 1, \dots, k$. Przyjęto standardowy nieobciążony estymator odchylenia standardowego postaci:

$$\hat{\sigma} = \sqrt{\frac{1}{k-1} \sum_{i=1}^k (y_i - \bar{y})^2},$$

gdzie \bar{y} oznacza średnią arytmetyczną przyrostów y_1, \dots, y_k .

W modelu wymagane jest również podanie parametru dryfu r dochodzącego po momencie zmiany procesu. Jest to parametr o tyle problematyczny, że ciężko go skalibrować na podstawie danych historycznych, skoro dotyczy dryfu, który wystąpi dopiero „w przyszłości”. Trzeba go zatem w pewnym sensie zgadnąć. Do jego wyznaczenia można jednak podejść w bardziej praktyczny sposób. Mianowicie, mając dany pewien instrument finansowy czy aktuarialny, którego wartość zależy od natężenia śmiertelności, można przeprowadzić dla niego analizę wrażliwości na ten właśnie parametr. W ten sposób można określić, przy jakim dryfie r natężenie śmiertelności zmieniłoby się na tyle istotnie, że w istotnym dla nas stopniu zmieniłaby się też cena rozważanego instrumentu, i ten dryf przyjąć do modelu.

W tym artykule nie rozważa się żadnego konkretnego instrumentu, dlatego też trzeba przyjąć pewne przykładowe wartości parametru r . Odniesiono więc te wartości do poziomu odchylenia σ i badano dwa warianty: bardziej „rygorystyczny” – z $r = 2 \cdot \sigma$, oraz „łagodniejszy” – z $r = \sigma$. Istotna jest też kwestia znaku tego dryfu: dla analizowanych w kolejnym rozdziale danych przyjęto dodatni, gdyż obserwowane natężenie śmiertelności, mimo że wciąż maleje, to wraz z upływem lat wydaje się maleć coraz wolniej.

Ostatnimi parametrami, które trzeba wyznaczyć, zanim przejdzie się do obliczeń, są: intensywność λ w rozkładzie momentu zmiany θ oraz dwa parametry, od których zależy rozpatrywane przez nas kryterium (3) lub (5), czyli c i π . Jeśli chodzi o λ , to podobnie jak z r będą badane dwie jego wartości zadane „z góry”. Odpowiada on za średni czas oczekiwania do wystąpienia momentu zmiany dryfu w modelu (który jest jego odwrotnością), arbitralnie zatem zbadane zostaną wartości $\lambda = 0,25$ i $\lambda = 0,1$. Co do wartości π , czyli prawdopodobieństwa wystąpienia zmiany już w chwili rozpoczęcia obserwacji, to również arbitralnie przyjęto $\pi = 0,1$ – tak aby procedura nie zatrzymywała się od razu zbyt często, ale też aby model był pełny, przyjęto wartość niezerową.

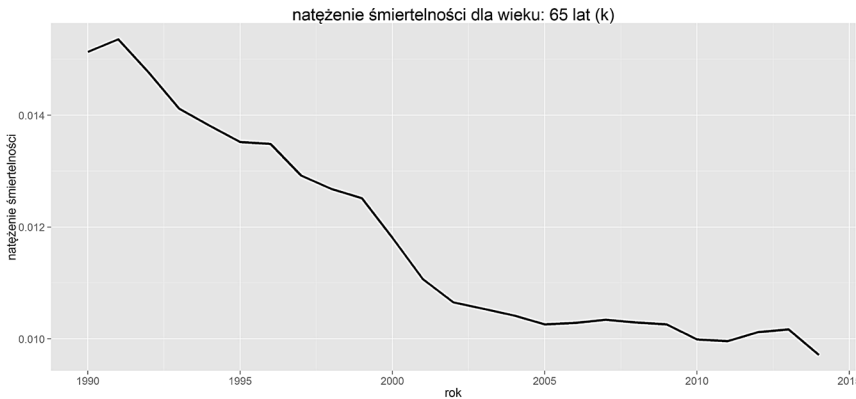
Na koniec ustalona zostanie wielkość c . Zwrócić należy uwagę, że odpowiada ona za stosunek wagi „fałszywego alarmu” do wagi średniego opóźnienia wykrycia momentu zmiany w rozważanym kryterium. Większe c spowoduje, że średnie opóźnienie będzie istotniejsze i częściej będziemy dostawali „fałszywe alarmy”, natomiast mniejsze skutkuje dłuższym średnim opóźnieniem wykrycia momentu zmiany. Należy przypomnieć, że wybrane kryterium uwzględnia prawdopodobieństwo wystąpienia „fałszywego alarmu”, a więc wartość tylko z przedziału $[0,1]$, oraz opóźnienie, które może przyjmować dowolnie duże nieujemne wartości. Zatem rozsądne (oraz poparte doświadczeniami empirycznymi) jest według autorów przyjęcie wartości $c = 0,5$. Mówiąc w uproszczeniu, można to rozumieć tak, że dwuletnie opóźnienie będzie miało ten sam „koszt” co wystąpienie „fałszywego alarmu”.

5. Detekcja zmiany dryfu natężenia śmiertelności populacji Polski w latach 1990-2014

Mając przedstawioną całą teorię, można przejść do analizy faktycznych danych. Wykorzystane zostaną polskie tablice trwania życia dostępnych w Internecie na stronie Głównego Urzędu Statystycznego

[GUS 2015], dotyczących lat 1990-2014. Jak już wspomniano wcześniej, uwaga zostanie skupiona na ustalonym wieku x lat i zbadaniu, jak zmieniało się dla niego natężenie śmiertelności w ciągu tych lat. Na początku dokładniej zostanie opisany jeden przypadek (zestaw parametrów), by dalej porównać go z innymi.

Najpierw autorzy przyjrzą się natężeniu śmiertelności kobiet w wieku 65 lat – obrazuje ją poniższy rys. 1.

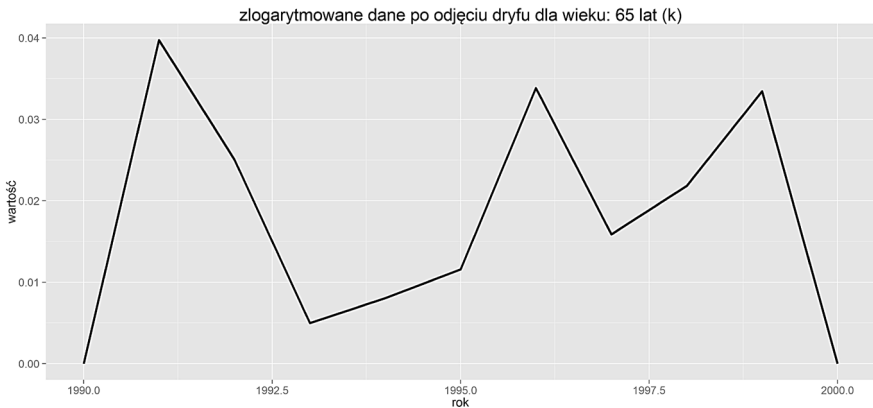


Rys. 1. Natężenie śmiertelności kobiet w wieku 65 lat, okres 1990-2014

Źródło: opracowanie własne.

Zgodnie z przypuszczeniami natężenie to ma tendencję malejącą – kobiety w Polsce w wieku 65 lat umierają stosunkowo coraz rzadziej – można jednak zauważyć, że w pierwszej dekadzie XXI wieku poziom natężenia śmiertelności wydaje się stabilizować, podczas gdy w latach 90. XX wieku wyraźnie się obniżał. Zanim uruchomiony zostanie algorytm na te dane, trzeba określić, które obserwacje posłużą do kalibracji modelu, a na których będzie już przeprowadzana procedura detekcji zmiany dryfu. Aby danych do kalibracji nie było zbyt mało, a zarazem żeby było ich wystarczająco w drugiej grupie – do detekcji, ustalono punkt podziału w roku 2000. Inaczej, zgodnie z oznaczeniami z poprzednich rozdziałów, za chwilę $t = 0$ przyjmuje się rok 2000. Dane z lat 90. posłużą do kalibracji modelu. Za jednostkę czasu (jeden krok algorytmu) przyjmuje się jeden rok.

Zgodnie z opisem w rozdziale dotyczącym kalibracji dane te najpierw zostaną zlogarytmowane, po czym odjęty zostanie od nich wyznaczony dryf i wyznaczone odchylenie standardowe przyrostów. Zlogarytmowane dane z lat 90. po odjęciu dryfu obrazuje rys. 2.



Rys. 2. Logarytm z natężenia śmiertelności kobiet w wieku 65 lat, po odjęciu dryfu – okres 1990-2000.

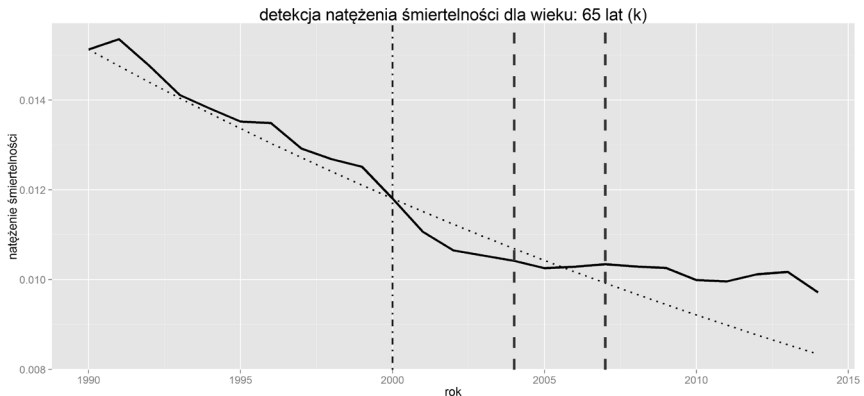
Źródło: opracowanie własne.

Odchylenie standardowe przyrostów powyższych danych wyniosło $\sigma = 0,022$.

Mając skalibrowany model, można przejść do detekcji. Zbadane zostaną jednocześnie przypadki $r = 1 \cdot \sigma$ oraz $r = 2 \cdot \sigma$. Dla pozostałych parametrów ustalono następujące wartości:

$$\lambda = 0,25, \quad c = 0,5, \quad \pi_0 = 0,1.$$

Na rysunku 3 czarną jednolitą linią oznaczono, tak jak poprzednio, historyczne natężenie śmiertelności w latach 1990-2014. Pionowa



Rys. 3. Detekcja zmiany dryfu natężenia śmiertelności kobiet w wieku 65 lat – okres 1990-2014

Źródło: opracowanie własne.

linia na poziomie roku 2000 (składająca się na zmianę z kresek i kropek) pokazuje chwilę $t=0$, czyli moment rozpoczęcia algorytmu. Kropkowana ukośna linia pokazuje sam dryf, skalibrowany zgodnie z danymi z lat 90. Wreszcie dwie grubsze pionowe, przerywane linie pokazują wykryte przez naszą procedurę detekcji momenty zmiany dryfu (dla większej wartości parametru r detekcja następuje później).

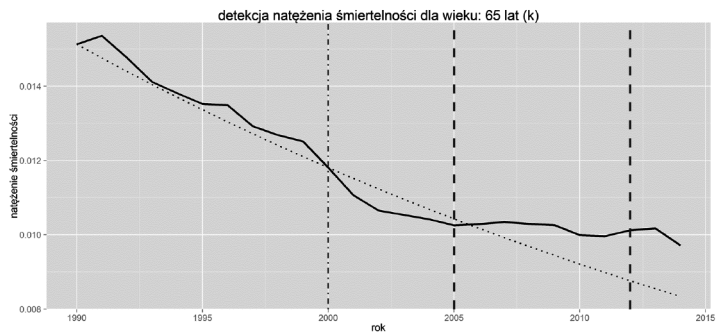
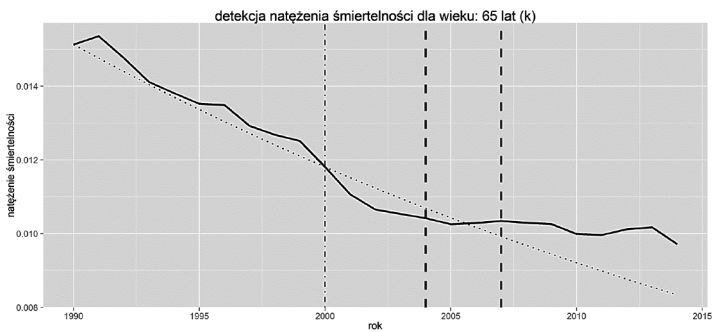
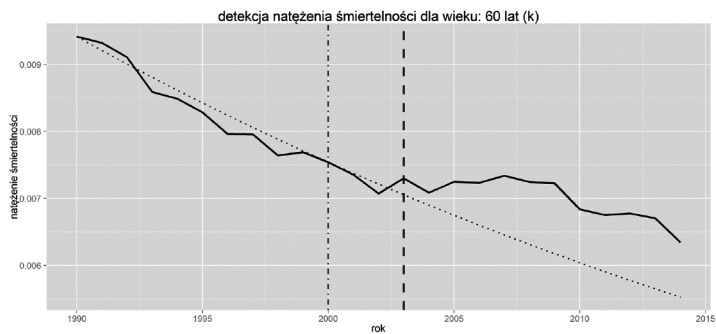
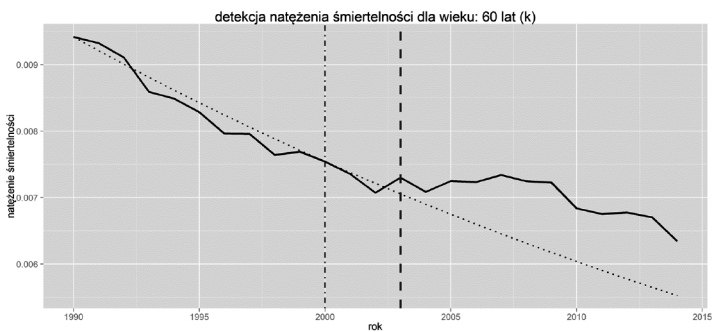
Jak widać, dla parametru $r=1 \cdot \sigma$ zmiana została wykryta w roku 2004, a dla $r=2 \cdot \sigma$ w roku 2007. W obu przypadkach algorytm wykrył zmianę, choć w różnych chwilach.

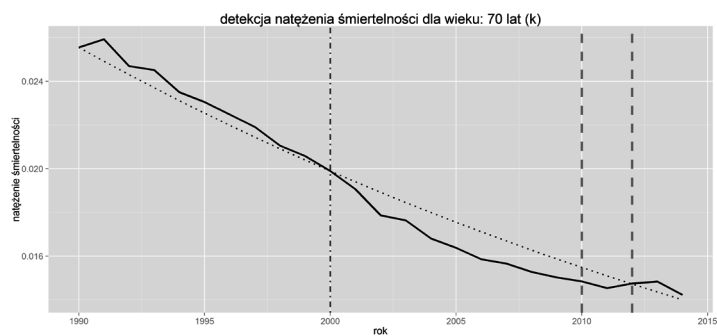
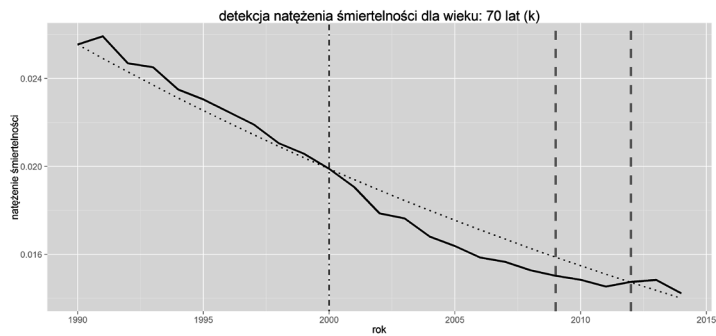
Zanim dokładniej opisane zostaną uzyskiwane wyniki oraz ich zależność od ustalonych parametrów, należy się przyjrzeć większej liczbie wykresów. Na rysunkach 4 i 5 oznaczenia są takie same jak powyżej i również badany jest dryf $r=1 \cdot \sigma$ i $r=2 \cdot \sigma$, gdzie zawsze dla większej wartości r detekcja występuje później bądź w tej samej chwili co dla mniejszej. Jeżeli wystąpiła w tej samej chwili, to jest tylko jedna linia przerywana. Jeżeli dla którejś (bądź obu) wartości r detekcja w ogóle nie nastąpiła, to przerywana linia narysowana jest na samym końcu wykresu (z prawej strony). Pierwsza kolumna odpowiada wartości parametru $\lambda = 0,25$, a druga $\lambda = 0,1$.

Na rysunku 5 przedstawiono podobne wykresy dla mężczyzn – dla wieku $x = 55, 60$ i 65 lat.

Na podstawie rys. 4 i 5 można wyciągnąć pewne wnioski. Pierwszą obserwacją jest fakt, że dla większej wartości parametru dryfu r detekcja następuje później niż dla mniejszej (bądź w tej samej chwili). Nie może nastąpić wcześniej, ponieważ statystyka oparta jest na ilorazie wiarygodności dwóch rozkładów normalnych o tej samej wariancji, zatem im bardziej przesunięty jest dryf, tym bardziej badane dane muszą odbiegać od rozkładu o średniej 0. Podobną zależność obserwujemy dla parametru λ : gdy jest mniejszy, to detekcja następuje później (lub nie wcześniej). To też nie dziwi – λ odpowiada za wartość średnią (jest jej odwrotnością) rozkładu *a priori* momentu zmiany θ , a nasza statystyka jest rozkładem *a posteriori*. Zmniejszając więc λ , opóźnia się średni moment wystąpienia zmiany, a więc zmniejsza prawdopodobieństwo, że zmiana już wystąpiła dla ustalonej chwili obserwacji t .

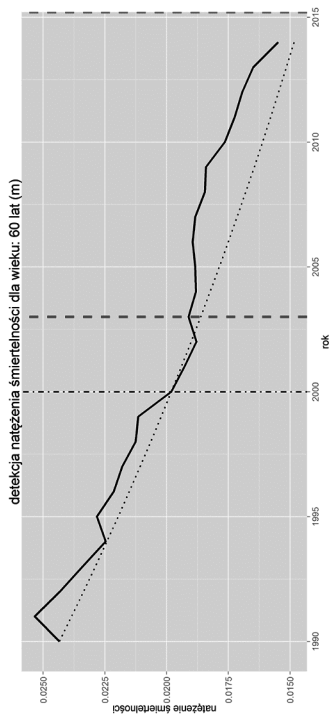
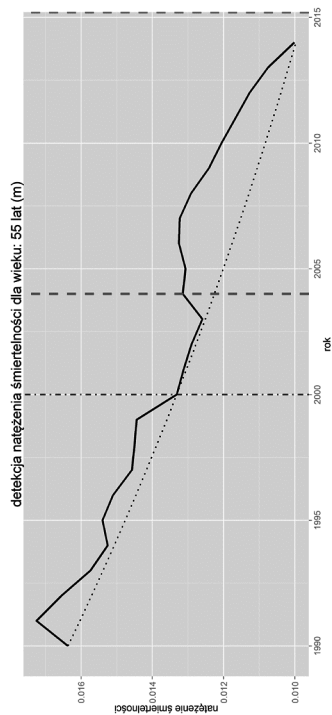
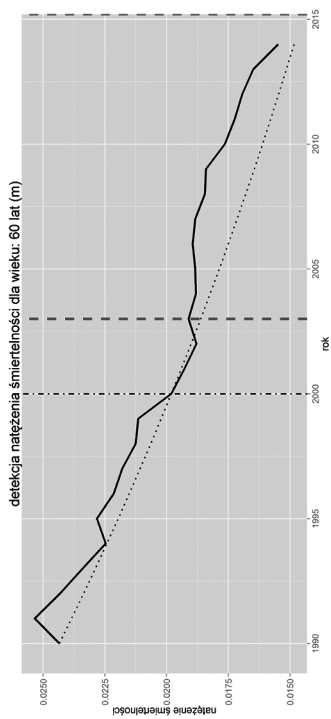
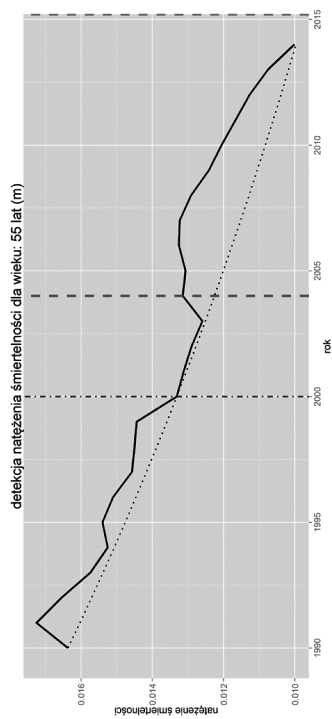
Zauważyć też należy, że różne wartości parametrów r i λ mogą dawać bardzo różne wyniki, ale mogą też dawać takie same. W przypadku kobiet w wieku 60 lat detekcja nastąpiła we wszystkich 4 przypadkach w tym samym 2003 roku. Z drugiej strony, dla kobiet w wieku 65 lat dla wszystkich 4 par tych parametrów detekcja wystąpiła

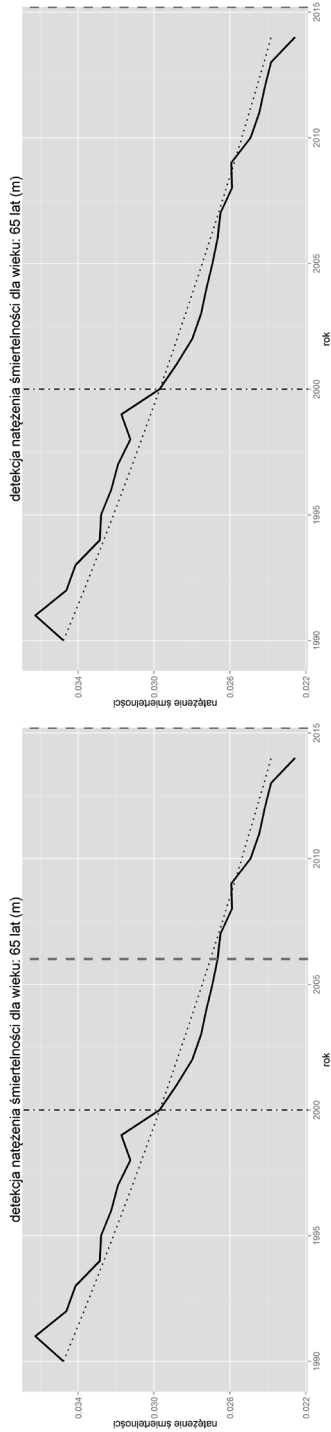




Rys. 4. Detekcja zmiany dryfu natężenia śmiertelności kobiet w Polsce w wieku 60, 65 i 70 lat, dla parametru $\lambda = 0,25$ (pierwsza kolumna) i $\lambda = 0,1$ (druga kolumna) – okres 1990-2014

Źródło: opracowanie własne.





Rys. 5. Detekcja zmiany dryfu natężenia śmiertelności mężczyzn w Polsce wieku 55, 60 i 65 lat dla parametru $\lambda = 0,25$ (pierwsza kolumna) i $\lambda = 0,1$ (druga kolumna) – okres 1990–2014

Źródło: opracowanie własne.

w różnych latach (ale zawsze wystąpiła). Wybór tych dwóch parametrów okazuje się jeszcze istotniejszy, gdy spojrzeć na wykresy dla mężczyzn. Dla wieku 55 oraz 60 lat na pierwszy rzut oka można powiedzieć, że około roku 2005 widać wyraźną zmianę dryfu, jednak algorytm ją wskazał jedynie w przypadku $r = 1 \cdot \sigma$. Dla $r = 2 \cdot \sigma$ algorytm nie wykrył zmiany dryfu na żadnym z 6 wykresów. Nie oznacza to jednak, że należy zawsze wybierać jak najmniejszą wartość parametru r – patrząc na wykres mężczyzn dla wieku 65 lat, raczej nie jest pożądane wskazanie zmiany dryfu, a dla $r = 1 \cdot \sigma$ procedura wykryła moment zmiany w roku 2006.

Można więc wyciągnąć wniosek, że parametry do algorytmu detekcji należy dostosować do potrzeb. Chcąc mieć procedurę, która wykryje choćby niewielką zmianę dryfu, ale za to akceptując częstsze „fałszywe alarmy”, przyjmując należy mniejsze wartości r i większe λ . Chcąc reagować dopiero na wyraźne zmiany i ustrzec się przed „fałszywymi alarmami”, należy postąpić odwrotnie.

6. Podsumowanie

Analiza przeprowadzona w artykule wskazuje, że algorytm optymalnej detekcji skonstruowany na podstawie uogólnionej wersji statystyki Shiryaeva-Roberts'a jest w stanie wskazać moment zmiany dryfu procesu. Zbadane dane pochodzące z polskich tablic trwania życia z lat 1990-2014 pokazują, że taka zmiana dryfu w procesie natężenia śmiertelności w niektórych przypadkach ustalonego wieku x lat wyraźnie występuje i opisywana procedura ją wykrywa – z różnym opóźnieniem zależnym od kalibracji parametrów. Jak pokazano, z doбором parametrów należy być ostrożnym również ze względu na możliwość wystąpienia „fałszywego alarmu”.

W ramach dalszych badań w tym obszarze autorzy planują przyjrzeć się też innym danym dotyczącym natężenia śmiertelności, czy to sięgającym więcej roczników wstecz, czy też dotyczącym innych krajów. Takie dane często są dużo bardziej nieregularne niż te zbadane w tym artykule, z widocznymi „skokami”, co powoduje, że nie nadają się do modelowania ruchem Browna. Dlatego też kolejnym krokiem w rozwoju procedury optymalnej detekcji jest znaczne uogólnienie samego modelu, tak aby uwzględniał skoki. Stawia to nowe wyzwania natury zarówno czysto matematycznej, jak i analitycznej.

Literatura

- El Karoui N., Loisel S., Salhi Y., 2015, *Minimax Optimality in Robust Detection of a Disorder Time in Poisson Rate*, <hal-01149749>, <https://hal.archives-ouvertes.fr/hal-01149749>.
- Gapeev P.V., 2005, *The disorder problem for compound Poisson processes with exponential jumps*, *The Annals Of Applied Probability*, vol. 15, s. 487-499.
- GUS, 2015, *Trwanie życia – tablice*, <http://www.stat.gov.pl> (04.06.2016).
- Peskir G., Shiryaev A.N., 2002, *Solving the Poisson Disorder Problem*, *Advances in Finance and Stochastics. Essays in Honour of Dieter Sondermann*, Springer-Verlag Berlin Heidelberg, New York, s. 295-312.
- Peskir G., Shiryaev A.N., 2006, *Optimal Stopping and Free-Boundary Problems. Lectures in Mathematics*, Birkhäuser, ETH Zürich.
- Shiryaev A.N., 1978, *Optimal Stopping Rules*, Springer-Verlag Berlin Heidelberg, New York.
- Shiryaev A.N., 2006, *From “Disorder” to Nonlinear Filtering and Martingale Theory*, *Mathematical Events of Twentieth Century*, Springer-Verlag Berlin Heidelberg i PHASIS Moscow, Niemcy, s. 371-397.

DRIFT CHANGE DETECTION IN MORTALITY RATE MODELS

Summary: Nowadays the insurance industry is facing huge challenges related to longevity risk, i.e. the risk that the trend of longevity growth significantly changes in the future. One of the crucial steps in dealing with it is identifying the change of the mortality rate drift observed in prospective life tables. The purpose of this article is to identify this change by casting the problem of quickest detection in the framework of optimal stopping theory. We construct generalized discrete-time Shiryaev’s-Roberts statistics and we use it in the analysis of Polish life tables from years 1990-2014. We model the logarithm of the mortal intensity by the Brownian motion that changes the zero drift into nonzero one at some random time. For this case we construct optimal stopping rule detecting substantial change of this drift. We also present calibration of above model using Central Statistical Office data and we carry out extensive statistical analysis showing huge potential of described statistics.

Keywords: mortality rate, detection, Shiryaev-Roberts statistics, Brownian motion, life tables.