

# Rekordy długowieczności i procesy Poissona

Wojciech NIEMIRO\*

## Część II: Dygresje i komentarze

Przypominamy podstawowe definicje i założenia opisywanego modelu.

$S(t)$  to z definicji prawdopodobieństwo dożycia  $t$  lat.

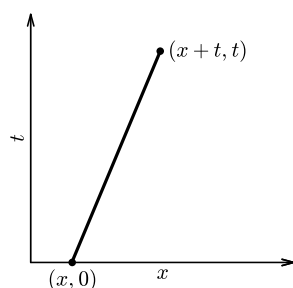
**T1** Każdy noworodek ma jednakową funkcję przeżycia  $S$ .

**T2** Długości życia różnych noworodków są statystycznie niezależne.

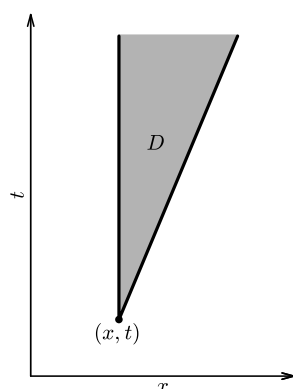
**N1** Dla dowolnego momentu  $x$ , prawdopodobieństwo urodzenia się dziecka w „krótkim” odcinku czasu  $(x, x + h]$  jest w przybliżeniu równe  $\lambda h$ , prawdopodobieństwo zaś urodzenia się więcej niż jednego dziecka jest tak małe, że możemy je zaniedbać.

**N2** Liczby noworodków pojawiających się w rozłącznych odcinkach czasu są statystycznie niezależne.

**TN** Długości życia wszystkich osobników są niezależne statystycznie od procesu narodzin.



Rys. 1. „Odcinek życia” osobnika, który urodził się w roku  $x$  i umarł w wieku  $t$  lat.



Rys. 2. Obszar  $D$  „zaczepiony” w punkcie  $(x, t)$ .

Przed miesiącem szukaliśmy prawdopodobieństwa tego, że umierający człowiek będzie starszy od wszystkich aktualnie żyjących. Zadanie wykonaliśmy.

Obliczyliśmy interesujące nas prawdopodobieństwo. Pozwólmy sobie teraz na kilka komentarzy i dygresji. Przypomnijmy najważniejszy wynik pomocniczy, który udowodniliśmy przed miesiącem.

**Stwierdzenie 1.** Prawdopodobieństwo tego, że osobnik umierający w wieku  $t$  lat jest starszy od wszystkich aktualnie żyjących, jest równe

$$\exp\{-\lambda R(t)\}, \quad \text{gdzie} \quad R(t) = \int_t^{\infty} S(u) du.$$

Warto przyjrzeć się bliżej funkcji  $R$ . Wyobraźmy sobie, że mamy „bardzo liczną” próbkę losową osobników o funkcji przeżycia  $S$ . Spodziewamy się, że spośród wszystkich  $n$  osobników liczba tych, którzy przeżyją przynajmniej  $t$  lat, jest bliska  $nS(t)$ . Tak jak w poprzednich rozważaniach, podzielimy przedział  $(t, \infty)$  na krótkie odcinki  $(t + ih, t + (i + 1)h]$ , gdzie  $i = 0, 1, \dots$ . Oczekujemy, że około  $n(S(t + ih) - S(t + (i + 1)h))$  osobników w naszej próbce będzie miało długość życia w  $i$ -tym przedziale, czyli w przybliżeniu  $t + ih$ . Jeśli teraz obliczymy średnią długość życia ponad  $t$  lat, to otrzymamy

$$\frac{1}{n} \sum_{i=0}^{\infty} ih n(S(t + ih) - S(t + (i + 1)h)) = \sum_{i=1}^{\infty} h S(t + ih) \approx R(t).$$

W rachunku prawdopodobieństwa nazywamy  $R(t)$  *wartością oczekiwaną* lub *średnią* czasu życia ponad  $t$  lat. Zauważmy, że mówimy tu o średniej pośród wszystkich osobników. Uznajemy, że ci, którzy wieku  $t$  nie dożyli, mają czas trwania życia ponad  $t$  lat równy 0. Jeśli obliczymy średnią pośród tylko tych osobników, którzy dożyli wieku  $t$ , to otrzymamy  $R(t)/S(t)$ . Poniższy wynik ma, z oczywistych względów, duże znaczenie dla ubezpieczeń życiowych.

**Stwierdzenie 2.** Oczekiwany dalszy czas życia osobnika, który dożył  $t$  lat, jest równy  $R(t)/S(t)$ .

W szczególności, średnia długość życia jest równa  $m = R(0)$ .

W realnych populacjach średnia długość życia jest skończona. Jednak z matematycznego punktu widzenia nic nie stoi na przeszkodzie, aby rozważać takie funkcje przeżycia  $S$ , że pole pod wykresem  $S$  jest nieskończone, czyli  $m = \infty$ . Każda taka nierosnąca funkcja  $S$ , że  $S(0) = 1$  i  $\lim_{t \rightarrow \infty} S(t) = 0$ , jest poprawną matematycznie funkcją przeżycia! W świecie, w którym obowiązują reguły T1, T2, N1, N2 i TN i jednocześnie  $m$  jest nieskończone, pewne zjawiska mogą się wydać paradoksalne.

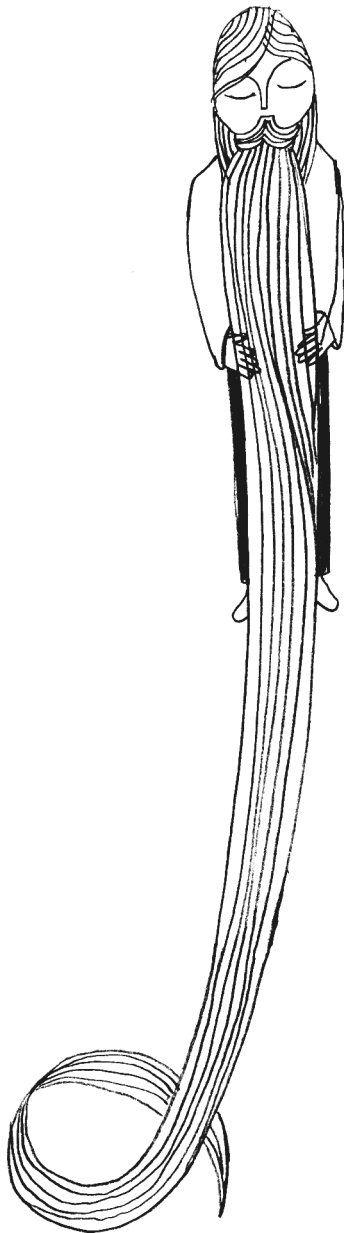
**Wniosek 1.** Jeśli średnia długość życia jest nieskończona, to w momencie śmierci każdego osobnika żyje nieskończenie wielu innych osobników od niego starszych.

Nikt nie jest nigdy najstarszy! Nieco dokładniej, teza Wniosku 1 jest zdarzeniem losowym, które zachodzi z prawdopodobieństwem równym 1.

**Dowód Wniosku 1.** Zdarzenie losowe, o którym mówi Stwierdzenie 1, zachodzi wtedy i tylko wtedy, gdy w obszarze  $D$  na rysunku 2 nie ma ani jednego końca „odcinka życia” (por. rys. 1). Jeśli  $m = \infty$ , to  $R(t) = \infty$  dla każdego  $t$ . Stosując formalnie wzór, widzimy, że w obszarze  $D$  nie ma „punktów śmierci” z prawdopodobieństwem  $\exp(-\lambda R(t)) = 0$  (można to uzasadnić nieco porządniej, modyfikując tylko nieznacznie podany poprzednio dowód).

Z prawdopodobieństwem 1, w obszarze  $D$  leży zatem przynajmniej jeden punkt. Nazwijmy teraz  $D$  „obszarem zaczepionym w punkcie  $(x, t)$ ” i zauważmy, że w nim jest zawarty każdy obszar zaczepiony w  $(x, t + i)$ , dla  $i = 1, 2, \dots$

\*Instytut Matematyki Stosowanej i Mechaniki, Uniwersytet Warszawski; Wydział Matematyki i Informatyki, Uniwersytet Mikołaja Kopernika, Toruń



W każdym z tych „mniejszych obszarów” leży przynajmniej jeden punkt, a zatem w obszarze  $\mathcal{D}$  leży nieskończenie wiele punktów, czego mieliśmy dowieść.

Z drugiej strony, w świecie, w którym obowiązują reguły T1, T2, N1, N2 i TN i jednocześnie  $m < \infty$ , też dzieją się rzeczy osobliwe.

**Wniosek 2.** *Jeśli średnia długość życia jest skończona, to w dowolnym, ustalonym momencie z niezerowym prawdopodobieństwem liczba aktualnie żyjących osobników jest równa zero.*

Co jakiś czas świat całkowicie się wyludnia. Na szczęście, nasze aksjomaty zapewniają, że dzieci będą rodzić się nieprzerwanie, z jednakową intensywnością.

**Zagadka 1.** *Jaki jest wzór na „niezerowe prawdopodobieństwo”, o którym mówi Wniosek 2?*

Dla ułatwienia dodajmy, że ten wzór wynika natychmiast z naszych poprzednich rozważań i zależy tylko od  $\lambda$  i  $m$ .

**Zagadka 2.** *Jaka jest średnia (oczekiwana) liczba osobników żyjących w ustalonym momencie?*

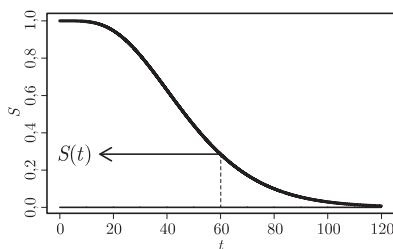
W tytule tego artykułu procesy Poissona występują w liczbie mnogiej. Założenia N mówią, że proces urodzeń jest jednorodnym procesem Poissona. Można wywnioskować, powołując się na Założenia T i TN, że proces zgonów jest też jednorodnym procesem Poissona z tą samą intensywnością  $\lambda$ . Żeby zobaczyć jeszcze jeden proces Poissona, zauważmy następujący fakt.

**Stwierdzenie 3.** *Zdarzenie polegające na tym, że w „krótkim” odcinku czasu  $(x, x + h]$  nastąpi śmierć osobnika, który narodził się w „krótkim” odcinku czasu  $(x - t, x - t + k]$ , ma prawdopodobieństwo w przybliżeniu równe  $\lambda\sigma(t)hk$ . Prawdopodobieństwo śmierci więcej niż jednego osobnika jest tak małe, że możemy je zaniedbać.*

Innymi słowy, możemy rozpatrywane zdarzenie wyrazić tak: w „małym równoległoboku” na płaszczyźnie znajdzie się jeden „punkt śmierci”. Widać tu analogię z warunkiem N1. Mały równoległobok można zastąpić, na przykład, „małym prostokątem”  $(x, x + h] \times (t, t + k]$ . Można udowodnić własność analogiczną do N2: liczby punktów w *rozłącznych obszarach* są statystycznie niezależne. Losowy zbiór „punktów śmierci” jest procesem Poissona na półpłaszczyźnie  $\{(x, t) : t > 0\}$ , z intensywnością  $\lambda\sigma(t)$ . Jest to proces jednorodny względem  $x$ , ale *niejednorodny* względem  $t$ , bo jego intensywność zależy od  $t$ . Niestety, nie możemy tu kontynuować opowieści o procesach Poissona. Czytelnik może sięgnąć do pięknej książki Kingmana *Procesy Poissona*, PWN, Warszawa, 2002.

\* \* \*

Na zakończenie popatrzmy na badane przez nas zjawisko z perspektywy statystycznej.

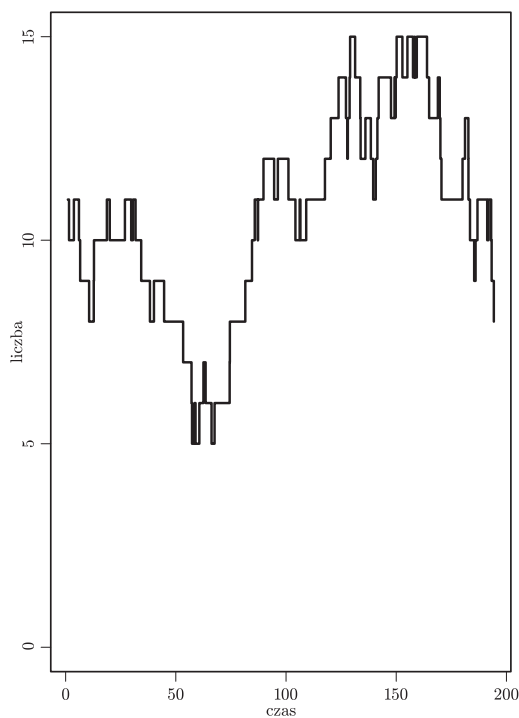


Rys. 3. Wykres funkcji przeżycia.

W tym celu udamy się na wyprawę do Symulandii. Jest to mały, wyspiarski kraj, którego mieszkańcy rodzą się i umierają ściśle według reguł T, N i TN. Funkcja przeżycia jest dokładnie taka, jak na rysunku 3 (znanym już z poprzedniej części artykułu), w szczególności  $m = 50$ , intensywność procesu urodzin jest równa  $\lambda = 0,2$ .

Co więcej, Główny Symulandzki Urząd Statystyczny dysponuje pełnymi danymi dotyczącymi osób, które urodziły się *lub* umarły w przedziale czasowym  $(0, 200]$  (według lokalnego kalendarza).

Dane dotyczą 82 osób i składają się z par  $(x_j, t_j)$ , gdzie  $x_j$  jest czasem urodzenia,  $t_j$  jest długością życia  $j$ -tej osoby ( $j = 1, \dots, 82$ ). Obliczyliśmy na ich podstawie kilka statystyk, które teraz przytoczymy. Są to „empiryczne odpowiedniki” wielkości, które analizowaliśmy.



Rys. 4. Liczba ludności Symulandii.

Wyprawę do Symulandii umożliwił pakiet statystyczny R.

R Development Core Team (2011). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

Każdy, kto ma komputer i dostęp do Internetu, może sobie zainstalować za darmo to potężne narzędzie obliczeniowe – i budować własne Symulandie. Zachęcam do takiej zabawy.

- W ciągu 200 lat zaobserwowano 44 urodzenia i 46 zgonów.
  - Porównaj z liczbą  $200 \cdot \lambda = 40$ .
- Średnia długość życia dla 82 osób była równa 48,72.
- 21 osób, czyli 25,61% spośród 82 przeżyło ponad 60 lat.
  - Porównaj z wartością funkcji przeżycia,  $S(60) = 25,51\%$ .
- Średnia długość życia ponad 60 lat, obliczona dla tych 21 osób, była równa 18,37 lat.
  - Porównaj z wielkością  $R(60)/S(60) = 18,17$ .
- W ciągu 200 lat ani razu liczba ludności Symulandii nie spadła do 0.
  - Porównaj z Zagadką 1.
- Średnia liczba ludności Symulandii w przedziale 200 lat była równa 10,31.
  - Porównaj z Zagadką 2 i z rysunkiem 4 (10,31 jest to nic innego, jak pole pod wykresem liczby ludności, podzielone przez 200).
- Na koniec statystyka dla nas najważniejsza. Spośród 46 zgonów 12 razy zdarzył się rekord długowieczności w sensie przez nas rozpatrywanym.
  - Porównaj ułamek  $12/46 = 26,09\%$  z liczbą 25,09% obliczoną zgodnie z uzyskanym przed miesiącem wzorem

$$\exp \left\{ -\lambda \int_t^{\infty} S(u) du \right\}.$$

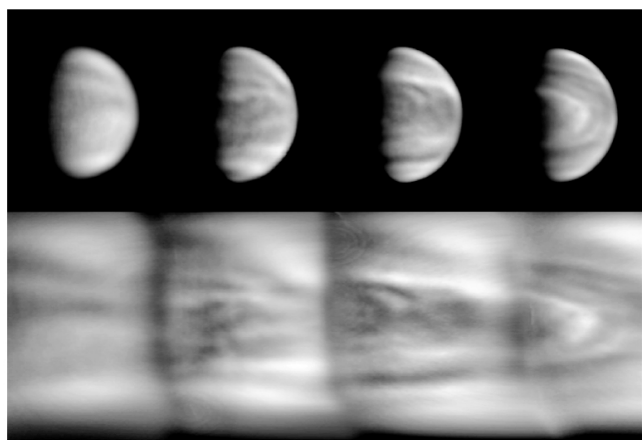
## Obserwacje naziemne Wenus w bazie danych ESA

Wenus, siostrzana planeta naszej Ziemi, jest bardzo wdzięcznym obiektem obserwacji astronomicznych, również tych amatorskich. W czerwcu 2012 roku byliśmy świadkami dość rzadkiego zjawiska astronomicznego, tzw. „przejścia”, czyli tranzytu Wenus na tle tarczy Słońca. W takich niecodziennych momentach oczy większości miłośników nieba zwracają się w stronę naszej Porannej Gwiazdy. Okazuje się jednak, że także na co dzień obserwacje Wenus mogą być nie tylko ciekawe, ale również przydatne naukowcom.

Przykładem wykorzystania amatorskich obserwacji Wenus przez profesjonalnych astronomów jest Venus Active Archive (VAA), czyli archiwum naziemnych zdjęć Wenus wykonanych

przez miłośników astronomii (<http://www.rssd.esa.int/vaa>). Powstało ono w celu uzupełnienia danych zbieranych przez Venus Express (VEX), czyli pierwszego satelitę zbudowanego przez Europejską Agencję Kosmiczną (ESA), który bada atmosferę Wenus. Niestety, nie jest on w stanie nieustannie monitorować nieba nad północną półkulą planety, ponieważ obiega ją po eliptycznej orbicie z apocentrum znajdującym się nad biegunem południowym. Położenie VEX na orbicie powoduje, że czasami części planety są widoczne z Ziemi, ale nie są widoczne z satelity. Dlatego właśnie dodatkowe obserwacje prowadzone z teleskopów naziemnych dostarczają brakujących i niezbędnych informacji na temat atmosfery Wenus.

VAA wykorzystuje możliwości amatorów astronomii, którzy uzyskują obrazy wykonane w podczerwieni lub ultrafiolecie za pomocą filtrów monochromatycznych oraz typowych kamer CCD. Zdjęcia te pozwalają, na przykład, na badanie jasności chmur Wenus względem całej jej powierzchni w zależności od czasu i mogą być użyte przez profesjonalnych astronomów w pracy naukowej. Przykładem mogą być obserwacje amatora astronomii, Berndta Gaehrkena, pokazane obok. Jest to zestawienie obrazów wykonanych pomiędzy 13 a 16 kwietnia 2007 roku za pomocą 80-cm teleskopu wyposażonego w kamerę internetową i filtr UV. Każdy obraz powstał ze złożenia około 4500 zdjęć. Mamy nadzieję, że pokazany przykład zachęci naszych Czytelników do obserwacji astronomicznych, które mogą mieć również wartość naukową.



April 13

April 14

April 15

April 16

Magdalena OTULAKOWSKA-HYPKA