

## PROBABILISTYCZNA IDENTYFIKACJA CZYNNIKA ZABURZAJĄCEGO LOSOWOŚĆ STRUMIENIA RUCHU STATKÓW WYCHODZĄCYCH Z PORTU SZCZECIN

*W artykule omówiona została koncepcja wyznaczania rozkładu prawdopodobieństwa czynnika regulującego wyjście statków z portu Szczecin. Wykorzystując rozkład czasu pomiędzy kolejnymi jednostkami opuszczającymi port w Szczecinie, wyznaczony na podstawie danych systemu VTS, oraz rozkład wykładniczy, odpowiadający losowemu strumieniowi ruchu statków, wyznaczono funkcje gęstości czasu przyspieszenia lub opóźnienia wyjścia statku, w stosunku do losowego strumienia ruchu. Wykorzystano w tym celu procedurę splotu funkcji gęstości zmiennych losowych.*

### WSTĘP

Badania parametrów ruchu statków na akwenach ograniczonych często wykazują [1, 2, 3], że jest on poddany działaniu różnych czynników zaburzających losowość zgłoszeń statków. Ma to również miejsce w przypadku ruchu statków wychodzących z portu. W pracach [1] i [2] autor wykazał, że strumień statków opuszczających port w Szczecinie nie ma charakteru poissonowskiego, czyli, że liczbę zgłoszeń statków w jednostce czasu nie można opisywać rozkładem Poissona, a co za tym idzie czas pomiędzy zgłoszeniami kolejnych jednostek torowych nie jest zgodny z rozkładem wykładniczym. Do opisu czasu pomiędzy trawersowaniem poszczególnych linii raportowych na torze Szczecin – Świnoujście wykorzystano mieszaniny rozkładów wykładniczych i jednostajnych [2]. Znając ten rozkład oraz teoretyczny rozkład wykładniczy, odpowiadający czasom pomiędzy zgłoszeniami statków w strumieniu zupełnie losowym, można wyznaczyć rozkład prawdopodobieństwa czasu przyspieszenia lub opóźnienia wyjścia statku, w stosunku do losowego strumienia ruchu. Możliwości zaawansowanych technicznie programów do obliczeń symbolicznych i numerycznych, jak np. *Mathematica*, pozwalają stosować procedury, których dawniej nie stosowano z powodu złożoności procesu obliczeniowego. Należą do nich również procedury związane z wyznaczaniem funkcji zmiennych losowych [3], zastosowane w niniejszej pracy.

### 1. ZMIENNE LOSOWE

Na odcinek czasowy pomiędzy wyjściem dwóch kolejnych statków z portu, składają się dwa elementy: jeden wynikający z charakteru losowego strumienia ruchu i drugi wynikający z charakteru obsługi statków w porcie. Zakładamy przy tym, że punkt raportujący statki, jest na tyle blisko portu, że na drodze pomiędzy nimi nie nastąpi wyprzedzenie jednego statku przez drugi (stąd wynika, że  $t$  jest większe od 0). Można więc czas  $t$  zapisać w następujący sposób:

$$t = t_1 + t_o \quad (1)$$

gdzie

$t$  – czas pomiędzy wyjściem dwóch kolejnych statków z portu,

$t_1$  – czas pomiędzy zgłoszeniami dwóch kolejnych statków w strumieniu losowym,

$t_o$  – czas przyspieszenia lub opóźnienia wyjścia statku, w stosunku do losowego strumienia ruchu.

W języku probabilistyki, zależność (1) można przedstawić następująco:

$$T = T_1 + T_o \quad (2)$$

gdzie

$T$  – zmienna losowa, opisująca czas pomiędzy dwoma kolejnymi statkami wychodzącymi z portu,

$T_1$  – zmienna losowa o rozkładzie wykładniczym, opisująca czas pomiędzy zgłoszeniami dwóch kolejnych statków w strumieniu losowym,

$T_o$  – zmienna losowa, opisująca czas przyspieszenia lub opóźnienia wyjścia statku, w stosunku do losowego strumienia ruchu.

Znając rozkłady zmiennych  $T$  oraz  $T_1$  i przy założeniu, że zmienne te są niezależne, można wyznaczyć rozkład prawdopodobieństwa zmiennej  $T_o$  jako różnicy zmiennych losowych ( $T - T_1$ ). W tym celu należy wykorzystać procedurę splotu funkcji zmiennych losowych [3]:

$$f(u) = \int_D f_1(u+v) \cdot f_2(v) dv \quad (3)$$

gdzie

$f_1$  – funkcja gęstości prawdopodobieństwa zmiennej  $T$ ,

$f_2$  – funkcja gęstości prawdopodobieństwa zmiennej  $T_1$ ,

$D$  – obszar całkowania, uzależniony od obszarów określoności funkcji  $f_1$  i  $f_2$ .

Obie funkcje gęstości  $f_1$  i  $f_2$  określone są tylko dla argumentów dodatnich, stąd obszar całkowania w procedurze (3) jest następujący:

$$f(u) = \begin{cases} \int_{-u}^{\infty} f_1(u+v) \cdot f_2(v) dv & \text{dla } u < 0 \\ \int_0^{\infty} f_1(u+v) \cdot f_2(v) dv & \text{dla } u \geq 0 \end{cases} \quad (4)$$

Wynika stąd, że zmienna losowa  $T_o$  może przyjmować wartości zarówno dodatnie jak i ujemne, dlatego zmienną tę nazwano czasem przyspieszenia lub opóźnienia wyjścia statku w stosunku do losowego strumienia ruchu.

## 2. TOR WODNY SZCZECIN-ŚWINOUJŚCIE

### 2.1. Punkt raportowy Inoujście

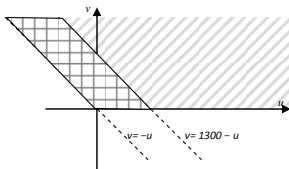
W przypadku toru wodnego Szczecin – Świnoujście, wykorzystano dane z systemu VTS. Jednym z punktów, w których rejestrowane są przejścia statków wychodzących z portu w Szczecinie jest punkt raportowy *Inoujście*. W drugiej połowie 2009 roku zanotowano w tym punkcie 1210 statków, opuszczających port Szczecin. Zmienna losowa, opisująca czas pomiędzy trawersowaniem linii raportowej, przez kolejne dwie jednostki ma rozkład o gęstości [2]:

$$f_1(x) = \begin{cases} 0,005157 e^{-0,0057x} + 0,000077 & \text{dla } 0 \leq x \leq 1300 \\ 0,005157 e^{-0,0057x} & \text{dla } x > 1300 \end{cases} \quad (5)$$

Natomiast zmienna  $T_1$  ma wtedy rozkład wykładniczy ze średnim czasem 222 minuty, czyli

$$f_2(x) = 0,0045 e^{-0,0045x} \text{ dla } x > 0 \quad (6)$$

W związku z tym, że funkcja  $f_1$  jest określona warunkowo, obszar całkowania procedury (3) należy podzielić na trzy części: dla  $u < 0$ , dla  $u$  z przedziału od 0 do 1300 i dla  $u > 1300$  (rysunek 1).



Rys. 1. Obszar całkowania procedury splotu

Otrzymujemy wtedy następującą całkę:

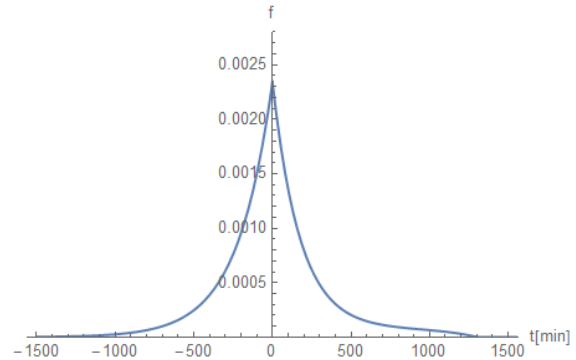
$$f(u) = \begin{cases} \int_{-u}^{1300-u} f_{11}(u+v) \cdot f_2(v) dv + \int_{1300-u}^{\infty} f_{12}(u+v) \cdot f_2(v) dv, & \text{dla } u < 0 \\ \int_0^{1300-u} f_{11}(u+v) \cdot f_2(v) dv + \int_{1300-u}^{\infty} f_{12}(u+v) \cdot f_2(v) dv, & \text{dla } u \in (0, 1300) \\ \int_0^{\infty} f_{12}(u+v) \cdot f_2(v) dv, & \text{dla } u > 1300 \end{cases} \quad (7)$$

gdzie  $f_{11}$  oznacza funkcję  $f_1$  określoną w przedziale od 0 do 1300, a  $f_{12}$  oznacza funkcję  $f_1$  określoną dla argumentów ponad 1300.

Po odpowiednim scałkowaniu uzyskano następującą funkcję gęstości zmiennej losowej  $T_o$ :

$$f(x) = \begin{cases} 0,0023 e^{0,0045x} & \text{dla } x < 0 \\ 0,000077 + 0,0023 e^{-0,00573x} - 0,0000002 e^{0,0045x} & \text{dla } x \in (0, 1300) \\ 0,0023 e^{-0,00573x} & \text{dla } x > 1300 \end{cases} \quad (8)$$

Wykres tej funkcji przedstawiono na rysunku 2. Wyznaczona funkcja gęstości pozwala obliczyć prawdopodobieństwo tego, że czas oczekiwania na zgłoszenie kolejnej jednostki opuszczającej port w Szczecinie zwiększy się lub zmniejszy w stosunku do czasu teoretycznego, opisanego rozkładem wykładniczym. Dodatnie wartości argumentów funkcji  $f$  oznaczają zwiększenie, a ujemne wartości zmniejszenie czasu oczekiwania na wyjście statku z portu.



Rys. 2. Wykres funkcji gęstości zmiennej  $T_o$

Bardziej szczegółowa analiza strumienia ruchu statków opuszczających port Szczecin wykazała dużą zmienność postaci rozkładu opisującego czas pomiędzy zgłoszeniami statków, w poszczególnych miesiącach 2009 roku. Analizy wykonano w innym niż poprzednio punkcie raportowym o nazwie Dok-5. Dane z całego analizowanego półrocza, podobnie jak w punkcie Inoujście, bardzo dobrze opisywała mieszanina rozkładów wykładniczego i jednostajnego. Natomiast w krótszych okresach lepsze dopasowanie dawały inne rozkłady.

### 2.2. Punkt raportowy Dok-5

Statystyczne testy zgodności chi-kwadrat Pearsona i Cramera von Misesa [5] wykazały, że oprócz mieszaniny rozkładów wykładniczego i jednostajnego, w niektórych okresach zdecydowanie lepsze dopasowanie dawały rozkłady Weibulla, Pareto, Gamma, czy też mieszaniny tych rozkładów z rozkładem jednostajnym. W niniejszym rozdziale zaprezentowane te, które zdecydowanie różniły się od rozkładu mieszanego, przedstawionego w rozdziale 2.1.

#### Strumień statków w miesiącu lipcu

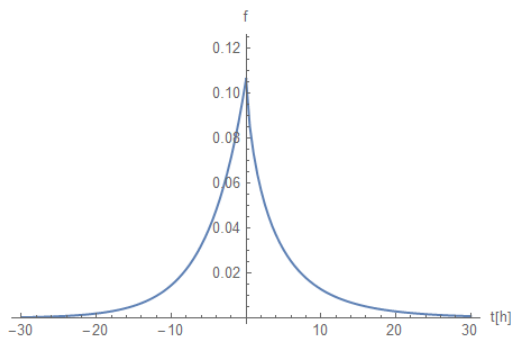
W tym okresie zarejestrowano 148 statków, dla których średni czas pomiędzy zgłoszeniami wynosił 5,01 godziny. W obu przeprowadzonych testach potwierdziła się hipoteza, że rozkład Weibulla z parametrami  $\alpha=0,809$  i  $\beta=4,574$ , bardzo dobrze opisuje czas pomiędzy trawersowaniem linii raportowej, przez kolejne dwie jednostki. Prawdopodobieństwa testowe (p-value) były bardzo wysokie: 0,59 w teście Cramera von Misesa i 0,68 w teście Pearsona. W związku z tym postać funkcji  $f_1$  z procedury (3) jest następująca:

$$f_1(x) = \frac{0,236 \exp(-0,292x^{0,809})}{x^{0,191}} \text{ dla } x > 0 \quad (9)$$

Natomiast zmienna  $T_1$  ma wtedy rozkład wykładniczy ze średnim czasem 5.01 h, czyli

$$f_2(x) = 0,199e^{-0,199x} \quad \text{dla } x > 0 \quad (10)$$

Zastosowanie formuły (4) okazało się zbyt złożone, by wyznaczyć analityczną postać funkcji gęstości zmiennej losowej  $T_o$ , dlatego poniżej przedstawiono tylko wykres numeryczny tej funkcji (rysunek 3).



Rys. 3. Numeryczna postać funkcji gęstości zmiennej  $T_o$

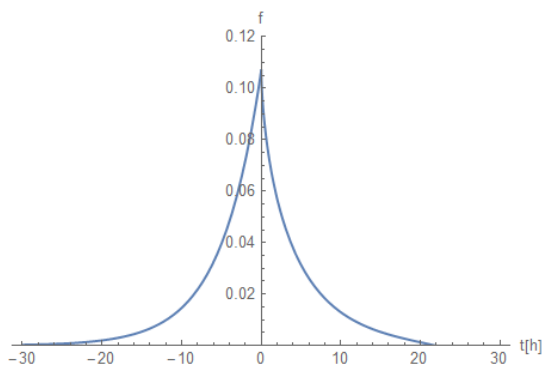
Równie dobre dopasowanie do danych z tego miesiąca daje rozkład Gamma z parametrami  $\alpha=0,72$  i  $\beta=6,9$ . Prawdopodobieństwa testowe również były bardzo wysokie: 0,49 w teście Cramera von Misesa i 0,61 w teście Pearsona. W związku z tym funkcja  $f_1$  z procedury (3) ma wtedy następującą postać:

$$f_1(x) = \frac{0,196 \exp(-0,145x)}{x^{0,28}} \quad \text{dla } x > 0 \quad (11)$$

Po zastosowaniu procedury (4) otrzymano:

$$f(x) = \begin{cases} 0,107 e^{0,199x} & \text{dla } x < 0 \\ 0,107 e^{0,199x} - 0,054 e^{-0,145x} x^{0,72} \cdot {}_1F_1(1; 1,72; 0,344x) & \text{dla } x \geq 0 \end{cases} \quad (12)$$

Symbol  ${}_1F_1$  oznacza konfluentną funkcję hipergeometryczną stosowaną w rozwiązaniach równań różniczkowych fizyki matematycznej [4]. Wykres funkcji (12) przedstawiono na rysunku 4.



Rys. 4. Wykres funkcji gęstości zmiennej  $T_o$  dla rozkładu Gamma

### Strumień statków w miesiącu sierpniu

W tym okresie zarejestrowano 171 statków, dla których średni czas pomiędzy zgłoszeniami wynosił 4,36 godziny. Testy statystyczne wykazały, że najlepiej czas pomiędzy trawersowaniem linii raportowej przez kolejne dwie jednostki opisuje mieszanina rozkładów jednostajnego i Gamma. Prawdopodobieństwa testowe były

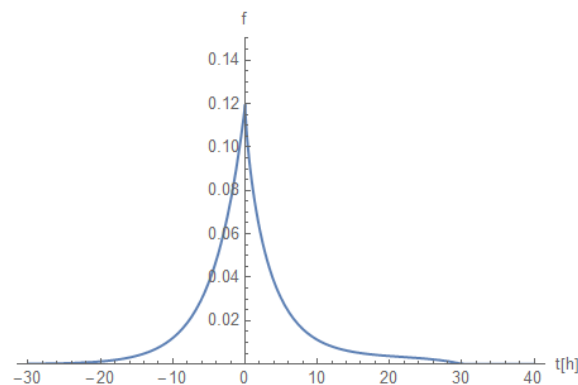
bardzo wysokie: 0,8 w teście Cramera von Misesa i 0,4 w teście Pearsona. W związku z tym postać funkcji  $f_1$  z procedury (3) jest następująca:

$$f_1(x) = \begin{cases} \frac{0,232 e^{-0,226x}}{x^{0,175}} + 0,0033 & \text{dla } 0 \leq x \leq 30 \\ \frac{0,232 e^{-0,226x}}{x^{0,175}} & \text{dla } x > 30 \end{cases} \quad (13)$$

W związku z tym, że funkcja  $f_1$  jest określona warunkowo, obszar całkowania, podobnie jak w punkcie 2.1, należy podzielić na trzy części: dla  $u < 0$ , dla  $u$  z przedziału od 0 do 30 i dla  $u > 30$ . Po odpowiednim scałkowaniu otrzymano funkcję gęstości zmiennej losowej  $T_o$ :

$$f(x) = \begin{cases} 0,119 e^{0,229x} & \text{dla } x < 0 \\ 0,0033 - 0,000003 e^{0,229x} + 0,115 e^{0,229x} + \\ -0,064 x^{0,825} \cdot {}_1F_1(0,825; 1,825; -0,455x) & \text{dla } x \in \langle 0, 30 \rangle \end{cases} \quad (14)$$

Wykres tej funkcji przedstawiono na rysunku 5.



Rys. 5. Wykres funkcji gęstości zmiennej  $T_o$  dla mieszaniny rozkładów Gamma i jednostajnego

### Strumień statków w miesiącu wrześniu

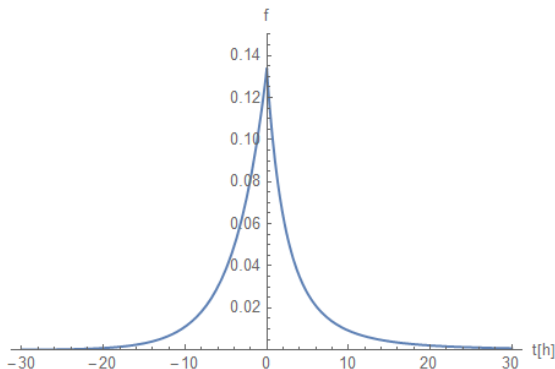
W tym okresie zarejestrowano 181 statków, dla których średni czas pomiędzy zgłoszeniami wynosił 4,02 godziny. W obu przeprowadzonych testach potwierdziła się hipoteza, że rozkład Pareto z parametrami  $k=7$ ,  $\alpha=2,5$  i  $\mu=0$ , bardzo dobrze opisuje czas pomiędzy trawersowaniem linii raportowej, przez kolejne dwie jednostki. Prawdopodobieństwa testowe były bardzo wysokie: 0,58 w teście Cramera von Misesa i 0,44 w teście Pearsona. W związku z tym postać funkcji  $f_1$  z procedury (3) jest następująca:

$$f_1(x) = \frac{324,105}{(7+x)^{3,5}} \quad \text{dla } x > 0 \quad (15)$$

Po zastosowaniu wzoru (4) otrzymano funkcję gęstości zmiennej losowej  $T_o$ :

$$f(x) = \begin{cases} 0,1336 e^{0,249x} & \text{dla } x < 0 \\ \frac{(-94,41 - 13,49x)e^{0,249x}}{7+x} + \frac{(226 + 32,3x) \cdot F_1(1; -1,5; 1,75 + 0,249x)}{(7+x)^{3,5}} & \text{dla } x \geq 0 \end{cases} \quad (16)$$

Wykres funkcji (16) przedstawiono na rysunku 6.



Rys. 6. Wykres funkcji gęstości zmiennej  $T_o$  dla rozkładu Pareto

## PODSUMOWANIE

Jak można zauważyć na prezentowanych rysunkach, wykresy wszystkich wyznaczonych funkcji gęstości zmiennej  $T_o$  są prawie identyczne, pomimo dużego zróżnicowania formuł je określających. Niezależnie czy do opisu strumienia ruchu statków wykorzystywano rozkłady Weibulla, Gamma, Pareto, wykładniczy czy mieszaniny tych rozkładów z rozkładem jednostajnym, to i tak rozkład prawdopodobieństwa czynnika zaburzającego losowość strumienia statków wychodzących z portu Szczecin, charakteryzował się podobnymi właściwościami. W rozkładzie tym można wyróżnić następujące cechy charakterystyczne:

- jest to rozkład z dominantą w punkcie 0,
- mediana jest ujemna, ale bliska 0,
- 99,9% wszystkich wartości mieści się w przedziale od -30 do 30 godzin (oznacza to, że większe odchylenie niż 30 godzin od losowego czasu oczekiwania na zgłoszenie kolejnej jednostki opuszczającej port Szczecin, jest praktycznie niemożliwe)
- rozkład jest leptokurtyczny (kurtoza dodatnia, wartości od 5,8 do 7,8)
- rozkład jest prawoskośny (współczynnik skośności dodatni, wartości od 1,8 do 2,3).

Zastosowanie zaawansowanych technik obliczeniowych w pakietach obliczeń symbolicznych, pozwoliło na zastosowanie procedury splotu funkcji gęstości różnych rozkładów. Wątpliwości może bu-

dzić złożoność formuł opisujących funkcję gęstości prawdopodobieństwa analizowanej zmiennej losowej  $T_o$  oraz zastosowanie funkcji specjalnych w prezentowanych gęstościach. Jednak możliwości obliczeniowe programu *Mathematica* są na tyle duże, że całkowanie prezentowanych formuł zajmowało ułamki sekund. Wyjątek stanowił przypadek z rozkładem Weibulla, gdzie udało się tylko uzyskać wartości numeryczne całek.

Formuły opisujące rozkład prawdopodobieństwa czynnika zaburzającego losowość strumienia statków wychodzących z portu Szczecin mogą zostać wykorzystane do budowy modeli symulujących strumień ruch statków opuszczających port Szczecin, a ogólna koncepcja może być wykorzystana w przypadku ruchu statków w innych portach. Możliwe jest też wykorzystanie wyznaczonego rozkładu do oceny pracy portu w różnych okresach lub porównywania różnych portów.

## BIBLIOGRAFIA

1. Kasyk L., *Analiza parametrów strumienia statków na torze wodnym Szczecin-Świnoujście*, „Logistyka” 2011, nr 3.
2. Kasyk L., *Non-Poisson character of vessel traffic on the Szczecin - Świnoujście fairway*, Zeszyty Naukowe, Akademia Morska w Szczecinie nr 32/2012.
3. Kasyk L., *Probabilistyczne metody modelowania parametrów strumienia ruchu statków na akwenach ograniczonych*. Wydawnictwo UTH, Radom 2012.
4. Kryszewski S. *Mechanika kwantowa*, Uniwersytet Gdański 2010.
5. Rogowski A., *Podstawy metod probabilistycznych w transporcie*, Wydawnictwo UTH w Radomiu, Radom 2012.

### Probabilistic identification of disturbing factor of random vessel traffic flow from the port of Szczecin

*In this paper a conception of determination of the probability distribution of the factor which modulates vessel traffic flow after departure from the port of Szczecin, has been presented. Using the time distribution between ships leaving port Szczecin, determined on the basis of VTS system, and theoretical exponential distribution of the time between reporting vessels in the random traffic flow, the probability density function of quickening or delaying ship departing from port, considering random vessel traffic flow, has been determined. To that end a procedure of convolution has been used.*

Autorzy:

dr hab. **Lech Kasyk**, profesor nadzwyczajny AM – Akademia Morska w Szczecinie, Instytut Technologii Morskich, l.kasyk@am.szczecin.pl