

## STATYSTYKA W PRAKTYCE

Piotr SULEWSKI

### Siatka prawdopodobieństwa uogólnionego rozkładu gamma

**Streszczenie.** *W badaniach statystycznych dużą popularność zyskują elastyczne rozkłady prawdopodobieństwa, których parametry są łatwe do oszacowania. W okresie poprzedzającym korzystanie z programów komputerowych zbudowanie siatek prawdopodobieństwa było możliwe tylko dla rozkładów o odwracalnej dystrybuancie, takich jak np. rozkład wykładniczy czy Weibulla. Dystrybuanta uogólnionego rozkładu gamma (URG) jest nie tylko nieodwracalna, lecz także nie ma formy analitycznej. Obecnie jednak, w dobie zaawansowanych możliwości informatycznych, dystrybuantę URG można odwrócić numerycznie przy pomocy różnych narzędzi, np. Microsoft Excel, Mathcad czy język R. Celem artykułu jest przedstawienie nowej metody tworzenia siatki prawdopodobieństwa URG wykorzystującej funkcję gęstości statystyki pozycyjnej oraz porównanie jej z metodami klasycznymi.*

**Słowa kluczowe:** siatka prawdopodobieństwa, uogólniony rozkład gamma, dystrybuanta empiryczna, metoda Monte Carlo.

**JEL:** C13, C15

---

W badaniach statystycznych istotną rolę odgrywają elastyczne rozkłady prawdopodobieństwa, których parametry są łatwe do oszacowania. Przykładem takiego elastycznego rozkładu jest z pewnością zaproponowany w roku 1962 przez Stacy'ego (1962) uogólniony rozkład gamma (URG), który znalazł także zastosowanie w teorii niezawodności jako rozkład czasu życia obiektów technicznych. Trzy lata później Stacy i Mihram (1965) zaproponowali technikę esty-

macji parametrów tego rozkładu za pomocą metody momentów (MM1). Niestety, metoda ta, choć prosta obliczeniowo, okazała się być bardzo nieefektywna (Sulewski, 2008). Konsekwencją tej małej efektywności jest tzw. bezpośrednia metoda momentów (MM2), która dla URG  $(a, b, c)$  polega na znalezieniu minimum funkcji celu (Sulewski, 2008).

$$M(x_i^*, a, b, c) = \sum_{k=1}^3 \left[ \sum_{i=1}^n (x_i^*)^k - a^k \cdot \Gamma\left(c + \frac{k}{b} \cdot \Gamma^{-1}(c)\right) \right]^2 \quad (1)$$

na podstawie wartości  $n$ -elementowej próby  $x_1^*, x_2^*, \dots, x_n^*$ . Za pomocą MM2 otrzymujemy estymatory zarówno bardziej efektywne, jak i mniej obciążone, co potwierdzają przebiegi funkcji gęstości estymatorów (wykr. 1). Aby jednak być w ocenie sprawiedliwym, należy stwierdzić, że MM1 i MM2 są metodami na miarę możliwości obliczeniowych czasów, w których się pojawiły. MM1 została opracowana w połowie lat sześćdziesiątych, gdy dostęp do komputerów był trudny.

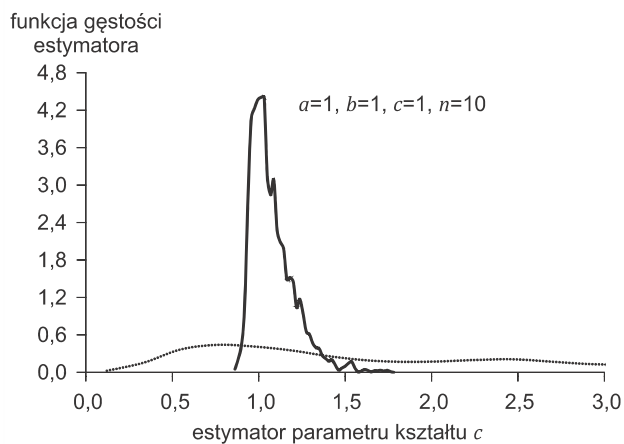
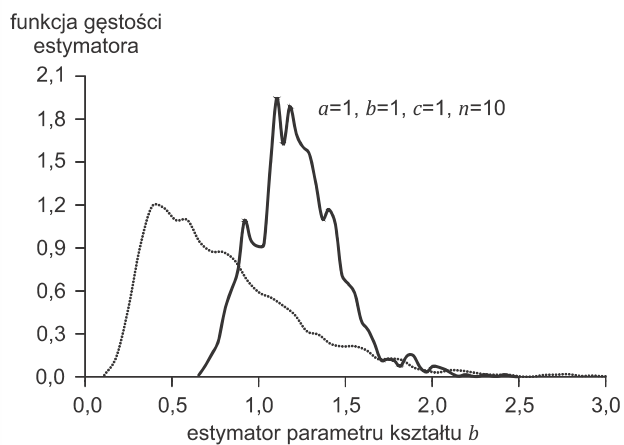
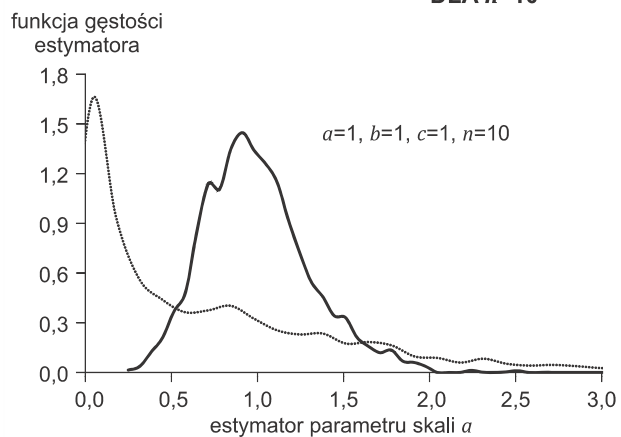
Siatka prawdopodobieństwa danego rozkładu jest prostokątnym układem współrzędnych z tak dobranymi skalami na osiach, aby wykreślona w tym układzie dystrybuanta teoretyczna była linią prostą. Jest ona narzędziem pozwalającym na sprawną ocenę zgodności rozpatrywanych danych doświadczalnych z rozkładem teoretycznym oraz na oszacowanie jego parametrów.

Poszukiwane są najczęściej takie rozkłady, dla których można zbudować siatkę prawdopodobieństwa, natomiast w okresie poprzedzającym komputery można było ją zbudować jedynie dla rozkładów o odwracalnej dystrybuancie. Dobrym przykładem jest rozkład wykładniczy i rozkład Weibulla. Rozkład normalny Gaussa jest wyjątkiem w tym względzie. Dystrybuanta URG jest nie tylko nieodwracalna, ale także nie ma formy analitycznej. Jednak obecnie — w dobie zaawansowanych komputerów — dystrybuantę URG można odwrócić numerycznie za pomocą metody Monte Carlo.

Z pewnością nie można kwestionować znaczenia siatek prawdopodobieństwa w zastosowaniach statystycznych. Siatkę prawdopodobieństwa skonstruowano m.in. dla takich rozkładów, jak: normalny (David, 1970; Sulewski, 2003), Weibulla (Drapella, Kosznik i Perzyńska, 1999) oraz Makehama (Kosznik, 1999). Inne pozycje literatury poświęcone temu zagadnieniu to m.in.: Wilk i Gnanadesikan (1968), Chambers (1983), Cleveland (1985) oraz Thode (2002).

W artykule przedstawiono główne charakterystyki URG oraz opisano generator liczb pseudolosowych dla tego rozkładu. Zaprezentowano trzy różne metody tworzenia siatki prawdopodobieństwa URG. Jakość estymacji parametrów URG porównano metodami symulacyjnymi. Wywód teoretyczny zakończono przykładem liczbowym i podsumowaniem.

Całą implementację komputerową wykonano w arkuszu Microsoft Excel z wykorzystaniem narzędzia Solver oraz edytora programistycznego Visual Basic for Applications (VBA).

**WYKR. 1. ROZKŁADY ESTYMATORÓW PARAMETRÓW URG UZYSKANE MM1 I MM2  
DLA  $n=10$** 

----- MM1  
————— MM2

### UOGÓLNIONY ROZKŁAD GAMMA

W zastosowaniach statystycznych rozkład powinien być elastyczny, aby możliwe było jego dopasowanie do danych. Zmienne rozkładu zwane parametrami są argumentami jego dystrybuanty. Parametry te — jeśli jest to niezbędne — umożliwiają uzyskiwanie różnych przebiegów dystrybuanty. Idealnym przykładem takiego elastycznego rozkładu jest URG.

W tej części pracy przedstawiono główne właściwości URG, mianowicie: funkcję gęstości, dystrybuantę oraz momenty. Funkcję gęstości oraz dystrybuantę URG opisano także w pracy Sulewskiego (2014), przy czym więcej informacji o tym rozkładzie można znaleźć we wcześniejszym opracowaniu (Sulewski, 2008).

URG jest rozkładem o takiej postaci analitycznej typu wykładniczo-potęgowego, która daje mu pożądaną elastyczność. Jego funkcja gęstości wyrażona jest wzorem (Stacy, 1962):

$$f(x; a, b, c) = \frac{b}{a\Gamma(c)} \left(\frac{x}{a}\right)^{bc-1} \exp\left[-\left(\frac{x}{a}\right)^b\right] \quad (x \geq 0) \quad (2)$$

gdzie:

$b > 0$ ,  $c > 0$  — parametry kształtu,  
 $a > 0$  — parametr skali.

URG zawiera w sobie szereg szczególnych przypadków — znanych z zastosowania w statystyce i teorii — niezawodności rozkładów (Sulewski, 2008). Rodzinę tych rozkładów prezentuje zestawienie. Prawdziwą zaletą rozkładu nie jest to, że inne rozkłady są jego szczególnymi przypadkami, lecz fakt, że wypełnia „przestrzeń” między nimi. Innymi słowy URG umożliwia opis parametryczny rozkładów empirycznych nieopisywalnych rozkładami tradycyjnymi, wyszczególnionymi w poniższym zestawieniu.

#### ZESTAWIENIE SZCZEGÓLNYCH PRZYPADKÓW URG

Rozkłady	Parametry URG		
	$a$	$b$	$c$
Wykładniczy .....	dowolny	1	1
Gamma .....	dowolny	1	dowolny
Erlanga .....	dowolny	1	całkowity
Weibulla .....	dowolny	dowolny	1
Pół-normalny .....	$\sigma\sqrt{2}$ ( $\sigma > 0$ )	1	0,5
Chi-kwadrat .....	2	2	$n/2$ ( $n \in N$ )
Chi .....	$\sqrt{2}$	2	$n/2$ ( $n \in N$ )
Rayleigha .....	$\sigma\sqrt{2}$ ( $\sigma > 0$ )	2	1
Maxwella-Boltzmannna .....	$\sigma\sqrt{2}$ ( $\sigma > 0$ )	2	1,5

U w a g a.  $n$  — liczba stopni swobody.

Ź r ó d ł o: opracowanie własne — modyfikacja tablicy z opracowania: Stacy i Mihram (1965, s. 351).

Wybrane zestawy wartości parametrów przedstawiono w tabl. 1, natomiast wykr. 2a prezentuje przebiegi funkcji gęstości URG dla tych zestawów.

**TABL. 1. PRZYKŁADOWE ZESTAWY  
WARTOŚCI PARAMETRÓW URG**

Zestaw	Parametry		
	skali $a$	kształtu $b$	kształtu $c$
1 .....	1	1	1
2 .....	1	2	3
3 .....	1	3	0,25
4 .....	1	0,9	0,9
5 .....	1	0,9	3

Ź r ó d ł o: opracowanie własne.

Korzystając z definicji dystrybuanty zmiennej losowej ciągłej dla URG mamy:

$$F(x; a, b, c) = \frac{b}{a\Gamma(c)} \int_0^x \left(\frac{t}{a}\right)^{bc-1} \exp\left[-\left(\frac{t}{a}\right)^b\right] dt \quad (3)$$

Wyrażenie (3) można zapisać w postaci bardziej dogodnej do obliczeń za pomocą niepełnej funkcji gamma (Kroese, Taimre i dan Botev, 2011):

$$\Gamma_n(c, u) = \int_0^u x^{c-1} \exp(-x) dx \quad (4)$$

Niech

$$u = \left(\frac{t}{a}\right)^b \quad (5)$$

wówczas w wyniku nieskomplikowanych przekształceń otrzymujemy:

$$\frac{t}{a} = u^{1/b} \Rightarrow dt = \frac{a}{b} u^{1/b-1} du \quad (6)$$

Podstawiając wzory (5) i (6) do (3) uzyskano:

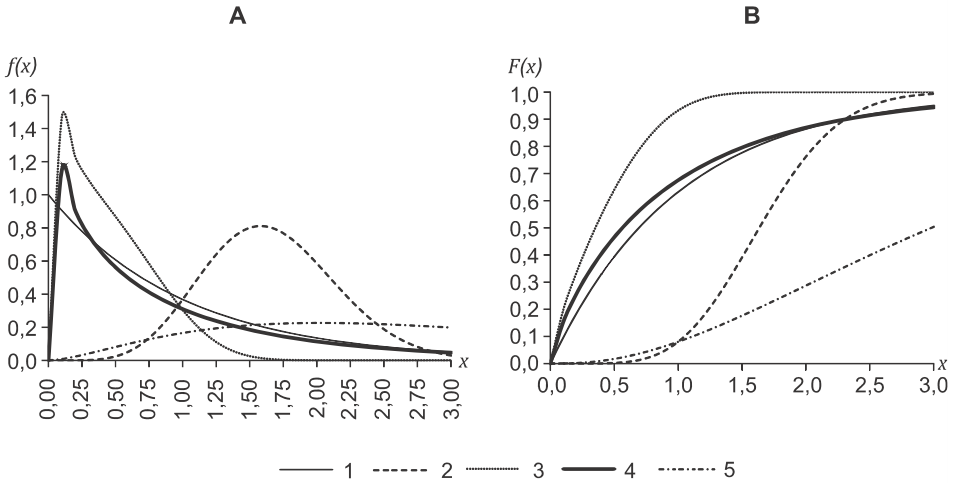
$$F(x; a, b, c) = \frac{1}{\Gamma(c)} \int_0^{(x/a)^b} u^{c-1} \exp(-u) du \quad (7)$$

Wyrażenie na dystrybuantę teoretyczną URG po uwzględnieniu wzoru (4) ma następującą postać:

$$F(x; a, b, c) = \frac{\Gamma_n[c, (x/a)^b]}{\Gamma(c)} \tag{8}$$

Dysponując procedurami numerycznymi na niepełną funkcję gamma wyznaczono wartości dystrybuanty URG ze wzoru (8). Przykładowe wykresy dystrybuanty dla zestawów wartości parametrów URG zaprezentowanych w tabl. 1 prezentuje wykr. 2b.

**WYKR. 2. FUNKCJA GĘSTOŚCI (A) I DYSTRYBUANTA (B) WYBRANYCH ROZKŁADÓW Z RODZINY URG**



Źródło: jak przy wykr. 1.

Z wykr. 2a wynika, że dobierając wartości parametrów URG można uzyskać przebiegi funkcji gęstości zarówno asymetrycznej, jak i symetrycznej.

Wyrażenie na momenty zwykłe URG, na mocy wyrażenia (2), ma następującą postać:

$$\alpha_k = \frac{b}{a \cdot \Gamma(c)} \int_0^\infty x^k \left(\frac{x}{a}\right)^{bc-1} \left[-\left(\frac{x}{a}\right)^b\right] dx \tag{9}$$

i postępując podobnie jak przy wyznaczaniu dystrybuanty URG otrzymujemy:

$$\alpha_k = a^k \cdot \frac{\Gamma\left(c + \frac{k}{b}\right)}{\Gamma(c)} \tag{10}$$

Wyrażenie na skośność URG, zwaną także współczynnikiem asymetrii, ma następującą postać:

$$\gamma_1 = \frac{\alpha_3 - 3\alpha_2\alpha_1 + 2\alpha_1^3}{(\alpha_2 - \alpha_1^2)^{3/2}} \quad (11)$$

natomiast kurtoza, zwana także współczynnikiem ostrości grzbietu, wyrażona jest wzorem:

$$\gamma_2 = \frac{\alpha_4 - 4\alpha_3\alpha_1 + 6\alpha_2\alpha_1^2 - 3\alpha_1^4}{(\alpha_3 - \alpha_1^2)^2} - 3 \quad (12)$$

W pracy Sulewskiego (2014) URG wykorzystano także do generowania wartości tablicy dwudzielczej  $2 \times k$ .

### GENERATORY LICZB PSEUDOLOSOWYCH O URG

Z zamieszczonego wcześniej zestawienia wynika, że rozkład gamma jest szczególnym przypadkiem rozkładu URG z parametrem  $b = 1$ . Zatem funkcja gęstości rozkładu gamma na podstawie wzoru (2) jest wyrażona następującym wzorem:

$$g(x; a, c) = \frac{1}{a\Gamma(c)} \left(\frac{x}{a}\right)^{c-1} \exp\left(-\frac{x}{a}\right) \quad (x \geq 0) \quad (13)$$

Niech  $z = (x/a)^b$  wówczas  $x = az^{1/b}$  oraz  $dx/dz = (a/b)z^{1/b-1}$ . Przechodząc od zmiennej losowej  $X$  o rozkładzie URG do nowej zmiennej losowej  $Z$ , we wzorze (2) otrzymano:

$$f(z; a, c) = \frac{a}{b} z^{1/b-1} \frac{b}{a\Gamma(c)} (z^{1/b})^{bc-1} \exp\left[-(z^{1/b})^b\right] \quad (z \geq 0)$$

a następnie, w wyniku prostych przekształceń, uzyskano wzór w postaci:

$$f(z; c) = \frac{1}{\Gamma(c)} z^{c-1} \exp[-z] \quad (z \geq 0) \quad (14)$$

Z wyrażień (13) i (14) wynika, że  $g(z; a = 1, c) = f(z; c)$ . Zatem zmienna losowa

$$Z = \left(\frac{X}{a}\right)^b \quad (15)$$

ma rozkład gamma z parametrem skali  $a = 1$  i parametrem kształtu  $c$  (Mihram i Stacy, 1965, s. 350).

Z zależności (15) wynika, że w celu uzyskania generatora realizacji zmiennej losowej  $X$  o URG wystarczy skonstruować generator realizacji zmiennej losowej

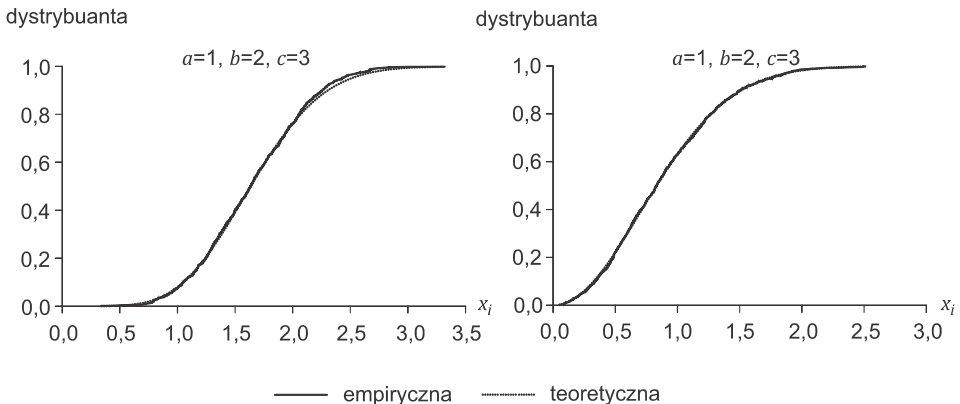
wej  $Z$  o rozkładzie gamma z parametrem skali  $a = 1$  i parametrem kształtu  $c$ . Funkcja użytkownika GenURG generująca realizacje zmiennej losowej  $X$  o URG powstała na podstawie następującego algorytmu (Zieliński, 1979, s. 98—102):

1. niech  $c - [c] = 0$  ( $c$  jest całkowite), wówczas:
  - 1.1.  $Z = -\ln(\prod_{i=1}^c U_i)$ , gdzie  $U_1, U_2, \dots, U_c$  są niezależnymi zmiennymi losowymi o rozkładzie  $U(0,1)$ ,
  - 1.2.  $X = aZ^{1/b}$ ;
2. niech  $c - [c] \neq 0$  ( $c$  nie jest całkowite), wówczas:
  - 2.1.  $Z_1 = -\ln(\prod_{i=1}^{[c]} U_i)$ ,
  - 2.2.  $Z_2 = -\ln(U)$ , gdzie  $U$  jest niezależną zmienną losową o rozkładzie  $U(0,1)$ ,
  - 2.3.  $d = c - [c]$ ,
  - 2.4.  $V_1 = U^{1/d}$ , gdzie  $V_1$  jest zmienną losową o rozkładzie potęgowym z parametrem  $d$ ,
  - 2.5.  $V_2 = U^{1/(1-d)}$ , gdzie  $V_2$  jest zmienną losową o rozkładzie potęgowym z parametrem  $1 - d$ ,
  - 2.6. jeżeli  $V_1 + V_2 > 1$ , wtedy następuje skok do punktu 2.4,
  - 2.7.  $Z_3 = V_1/(V_1 + V_2)$ ,
  - 2.8.  $Z = Z_1 + Z_2 Z_3$ ,
  - 2.9.  $X = aZ^{1/b}$ .

Dla  $n = 10^3$  realizacji zmiennej losowej URG otrzymanych za pomocą powyższego algorytmu i uporządkowanych rosnąco, obliczono wartości dystrybuanty teoretycznej URG (8) oraz wartości dystrybuanty empirycznej (wykr. 3):

$$F = \frac{i}{n+1} \quad (i = 1, 2, \dots, n) \quad (16)$$

### WYKR. 3. PRZEBIEGI DYSTRYBUANT EMPIRYCZNEJ I TEORETYCZNEJ URG





Wykr. 3 pokazuje, że przebiegi dystrybuant (8) i (16) pokrywają się, co potwierdza prawidłowe działanie generatora.

### SIATKA PRAWDOPODOBIENSTWA URG

Ta część pracy jest poświęcona trzem metodom tworzenia siatek prawdopodobieństwa URG.

#### Metoda I

Niech  $x_{(i)}^*$  będzie  $i$ -tą statystyką pozycyjną w  $n$ -elementowej próbie oraz niech  $x_i^0$  będą pierwiastkami następującego równania:

$$F(x_i^0; a^*, b^*, c^*) = \frac{i}{n+1} \quad i = 1, 2, \dots, n \quad (17)$$

gdzie:

$F$  — dystrybuanta teoretyczna URG,  
 $a^*, b^*, c^*$  — wartości parametrów tego rozkładu.

Dystrybuantę URG można wyznaczyć ze wzoru (8), a następnie numerycznie rozwiązać równanie nieliniowe względem  $x_i^0$ . Wartości tych pierwiastków są zbliżone do oczekiwanych wartości statystyk pozycyjnych. Siatka prawdopodobieństwa jest prostokątnym układem współrzędnych kartezjańskich. Empiryczne wykresy punktów w tym układzie mają rzędne równe  $x_i^0$  i odcięte równe  $x_{(i)}^*$ . W przypadku kiedy zbiór punktów rozciąga się wzdłuż linii prostej, wówczas można stwierdzić, że dane mają URG. W celu uzyskania takiego efektu należy tak dobrać wartości parametrów  $a^*, b^*, c^*$ , aby minimalizowały funkcję celu:

$$M_1(a^*, b^*, c^*) = \sum_{i=1}^n [x_{(i)}^* - x_i^0]^2 \quad (18)$$

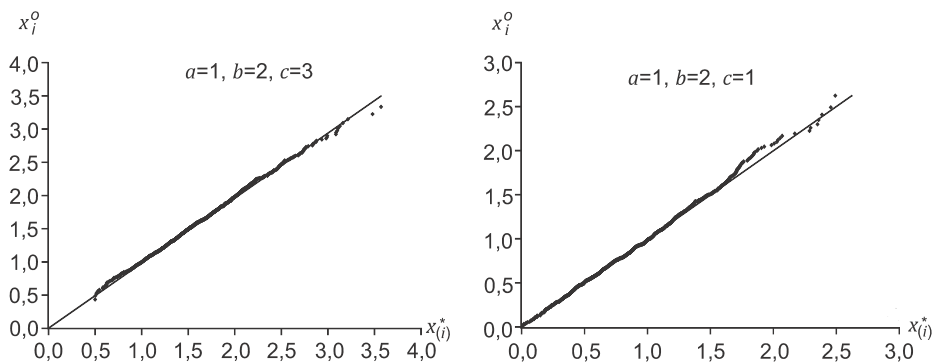
gdzie  $x_i^0$  jest rozwiązaniem równania (17).

Dane do wykr. 4 pochodzą z generatora URG dla wybranych wartości parametrów  $a, b, c$  i  $n = 1000$  oraz z zastosowania metody bisekcji do numerycznego rozwiązania równania (17).

Z wykr. 4 wynika, że siatka URG składa się z linii prostej nachylonej do osi odciętej pod kątem  $45^\circ$  oraz ze zbioru punktów, które pod wpływem odpowiednich zmian wartości parametrów  $a^*, b^*$  i  $c^*$  zbiegają ku tej prostej. Jest to prze-

ciwieństwo klasycznej metody najmniejszych kwadratów, w której linia prosta jest przesuwana do zbioru punktów.

WYKR. 4. SIATKA PRAWDOPODOBIEŃSTWA URG UZYSKANA METODĄ I



Źródło: jak przy wykry. 1.

## Metoda II

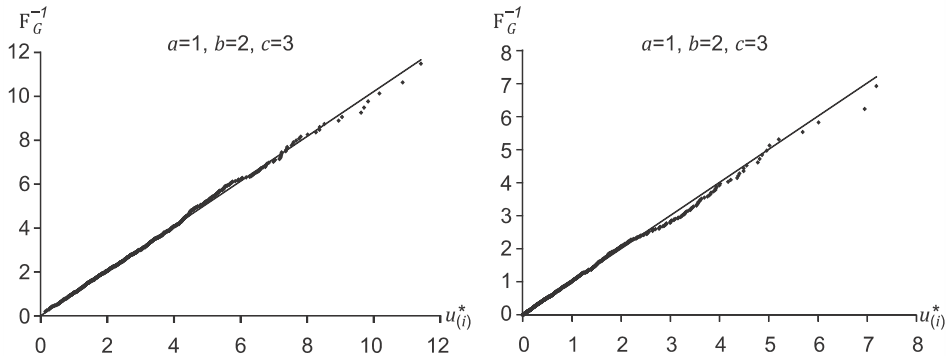
W dobie komputerów, jeżeli dysponuje się procedurą liczącą funkcję odwrotną do rozkładu gamma, który jest szczególnym przypadkiem URG dla  $b = 1$ , można skorzystać z formuły:

$$u_{(i)}^* = \left( \frac{x_{(i)}^*}{a^*} \right)^{b^*} \quad (19)$$

Wartości  $u_{(i)}^*$  ( $i = 1, \dots, n$ ) w równaniu (19) podlegające rozkładowi gamma należy nanieść na oś poziomą siatki prawdopodobieństwa, natomiast wielkości  $F_G^{-1}[i/(n+1), c^*]$  na oś pionową siatki, gdzie  $F_G^{-1}$  jest funkcją odwrotną do dystrybucyjną rozkładu gamma, czyli funkcją kwantylową tego rozkładu. Aby zbiór punktów rozciągał się wzdłuż linii prostej, należy tak dobrać wartości  $a^*$ ,  $b^*$  i  $c^*$ , aby minimalizowały funkcję celu:

$$M_2(a^*, b^*, c^*) = \sum_{i=1}^n \left[ \left( \frac{x_{(i)}^*}{a^*} \right)^{b^*} - F_G^{-1} \left( \frac{i}{n+1}, c^* \right) \right]^2 \quad (20)$$

Dane do wykry. 5 pochodzą z generatora URG dla wybranych wartości parametrów  $a$ ,  $b$ ,  $c$  i  $n = 1000$  oraz z zastosowania funkcji arkuszowej ROZKŁAD GAMMA.ODW( $i/(n+1)$ ;  $c^*$ ; 1).

**WYKR. 5. SIATKA PRAWDOPODOBIENSTWA URG UZYSKANA METODĄ II**

Źródło: jak przy wykr. 1.

Wnioski wynikające z wykr. 5 potwierdzają wnioski sformułowane na podstawie wykr. 4.

**Metoda III**

Wydaje się, że jest uzasadnione rozpatrywanie rozbieżności wartości zaobserwowanej statystyki pozycyjnej w próbie z teoretyczną najbardziej prawdopodobną wartością tej statystyki. Zatem metoda III korzysta z wartości modalnej funkcji gęstości statystyk pozycyjnych. Niech  $x_{(i)}^*$  będzie  $i$ -tą statystyką pozycyjną w  $n$ -elementowej próbie. Funkcja gęstości  $i$ -tej statystyki pozycyjnej ma następującą postać:

$$f_{i,n}(x; \oplus) = \frac{n!}{(i-1)!(n-i)!} f(x; \oplus) F(x; \oplus)^{i-1} [1 - F(x; \oplus)]^{n-i} \quad (21)$$

gdzie  $\oplus = a^*, b^*, c^*$  oraz  $f(x)$  i  $F(x)$  są dane odpowiednio wzorami (2) i (8).

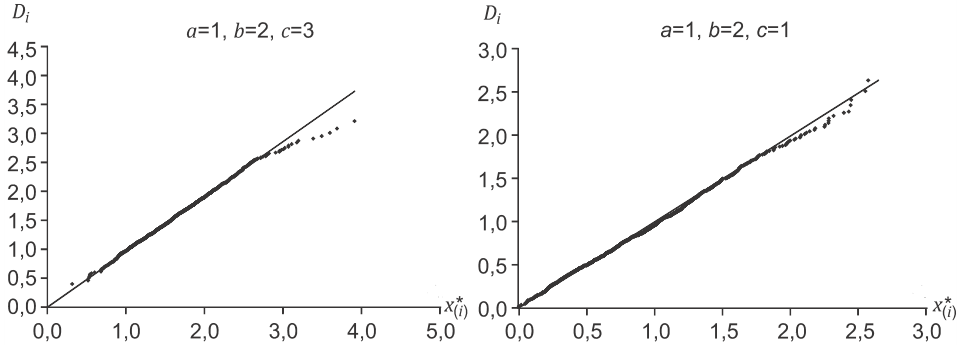
Siatka prawdopodobieństwa powstaje przez naniesienie wartości  $x_{(i)}^*$  na oś poziomą kartezjańskiego układu współrzędnych oraz wartości modalnych funkcji gęstości (21) na oś pionową. Aby zbiór punktów rozciągał się wzdłuż linii prostej, należy dobrać wartości  $a^*, b^*$  i  $c^*$  minimalizujące funkcję celu:

$$M_3(a^*, b^*, c^*) = \sum_{i=1}^n [x_{(i)}^* - D_i(a^*, b^*, c^*)]^2 \quad (22)$$

gdzie  $D_i(a^*, b^*, c^*)$  są wartościami modalnymi funkcji gęstości (21).

Dane do wykr. 6 pochodzą z generatora URG dla wybranych wartości parametrów  $a$ ,  $b$ ,  $c$  i  $n = 1000$  oraz z zastosowania metody bisekcji do znalezienia wartości modalnych funkcji gęstości (21).

WYKR. 6. SIATKA PRAWDOPODOBIEŃSTWA URG UZYSKANA METODĄ III



Źródło: jak przy wykr. 1.

Wnioski wynikające z wykr. 6 są identyczne jak z wykr. 4.

### EFEKTYWNOŚĆ ESTYMACJI

Wybór zestawów wartości parametrów przedstawiony w tabl. 2 nie był przypadkowy. Dobrano go w taki sposób, aby uzyskać wszystkie możliwe przebiegi funkcji ryzyka (funkcji intensywności uszkodzeń), będącej ilorzem funkcji gęstości  $f(x)$  i funkcji niezawodności  $R(x)$ , danej w postaci:

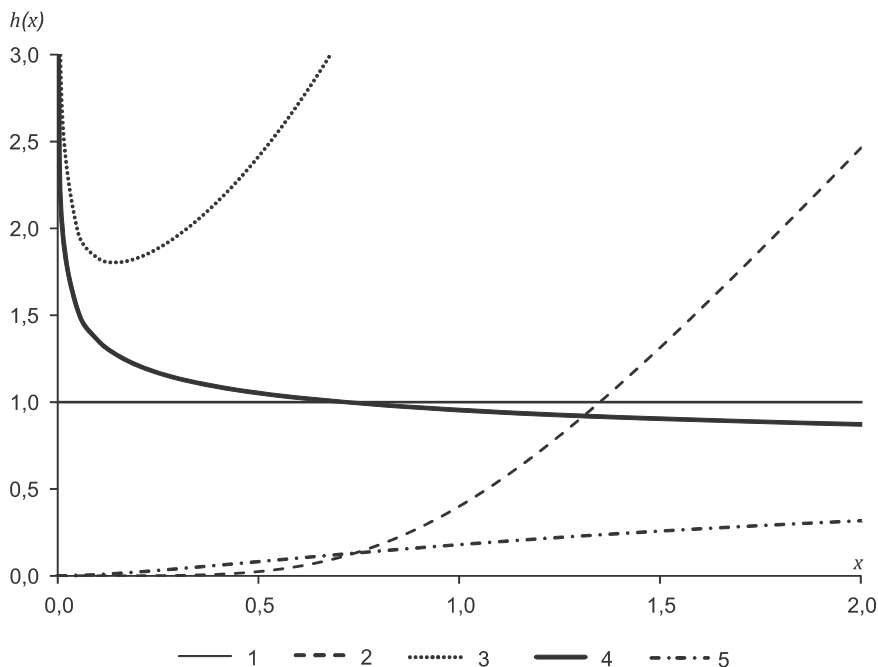
$$h(x) = \frac{f(x)}{R(x)} = \frac{f(x)}{1 - F(x)} \quad (23)$$

Funkcja ryzyka (23) jest istotną charakterystyką teorii niezawodności. Wspomniane przebiegi (wykr. 7) to: stały (zestaw 1), rosnący (zestaw 2), siodłowy (zestaw 3), malejący (zestaw 4) oraz grzbietowy (zestaw 5).

Wykr. 7 potwierdza, że dobierając wartości parametrów URG (tabl. 2), można uzyskać wszystkie możliwe przebiegi funkcji ryzyka tego rozkładu.

Korzystając z generatora liczb losowych URG porównano jakość estymacji parametrów tego rozkładu za pomocą siatek prawdopodobieństwa utworzonych metodami I—III na podstawie tej samej  $n$ -elementowej próby oraz dla każdego zestawu wartości parametrów (tabl. 2).

WYKR. 7. PRZEBIEGI FUNKCJI RYZYKA URG  
DLA ZESTAWÓW WARTOŚCI PARAMETRÓW Z TABL. 1



Źródło: jak przy wykresie 1.

Porównanie to dla każdej metody tworzenia siatki przeprowadzono za pomocą następującego algorytmu:

1. niech  $S_D = 0$ ,  $S_a = 0$ ,  $S_b = 0$ ,  $S_c = 0$ ;
2. kroki 2.1—2.8 powtórz  $N$  razy:
  - 2.1. ustal liczebność próby  $n$  oraz wartości parametrów  $a$ ,  $b$  i  $c$ ,
  - 2.2. wygeneruj  $n$ -elementową próbę mającą URG,
  - 2.3. utwórz siatki prawdopodobieństwa,
  - 2.4. wyznacz wartość funkcji celu  $M$ ,
  - 2.5. wyznacz wartości  $a^*$ ,  $b^*$  i  $c^*$ , dla których funkcja celu  $M$  osiąga wartość najmniejszą,
  - 2.6. oblicz  $S_a = S_a + |a^* - a|$ ,  $S_b = S_b + |b^* - b|$ ,  $S_c = S_c + |c^* - c|$ ,
  - 2.7. oblicz odległość euklidesową  $D$  między punktami  $A(a, b, c)$  i  $B(a^*, b^*, c^*)$

$$D = |AB| = \sqrt{(a - a^*)^2 + (b - b^*)^2 + (c - c^*)^2}, \quad (24)$$

- 2.8. oblicz  $S_D = S_D + D$ ;
3. oblicz  $\bar{S}_D = S_D/N$ ,  $\bar{a} - a^* = S_a/N$ ,  $\bar{b} - b^* = S_b/N$ ,  $\bar{c} - c^* = S_c/N$ .

W tabl. 2 przedstawiono wartości średnie odległości euklidesowej  $\bar{S}_D$  oraz wartości wyrażeń  $\overline{a - a^*}$ ,  $\overline{b - b^*}$ ,  $\overline{c - c^*}$  dla różnych zestawów wartości parametrów URG (tabl. 1), gdy  $N = 10^4$  oraz liczebność próby  $n = 100$ .

**TABL. 2. EFEKTYWNOŚĆ ESTYMACJI PARAMETRÓW URG**

Zestawy	Metody	$\bar{S}_D$	$\overline{a - a^*}$	$\overline{b - b^*}$	$\overline{c - c^*}$
1	I	0,146	0,078	0,081	0,069
	II	0,556	0,433	0,145	0,292
	III	0,041	0,022	0,019	0,023
2	I	0,349	0,082	0,154	0,280
	II	0,472	0,102	0,166	0,410
	III	0,059	0,011	0,051	0,010
3	I	0,386	0,060	0,371	0,030
	II	0,435	0,167	0,359	0,092
	III	0,119	0,030	0,113	0,004
4	I	0,115	0,075	0,119	0,117
	II	0,336	0,172	0,126	0,209
	III	0,070	0,017	0,110	0,050
5	I	0,219	0,057	0,039	0,197
	II	0,501	0,236	0,072	0,414
	III	0,099	0,012	0,014	0,074

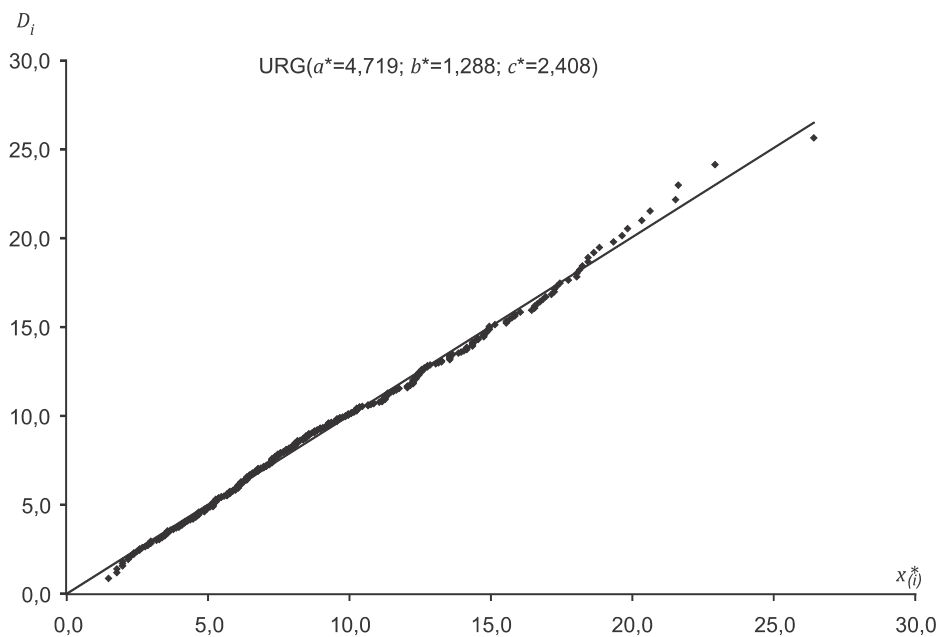
Źródło: jak przy tabl. 1.

Z tabl. 2 wynika, że proponowana metoda III charakteryzuje się najlepszą efektywnością estymacji ze względu na średnią odległość uzyskanych oszacowań  $a^*$ ,  $b^*$  i  $c^*$  od wartości teoretycznych  $a$ ,  $b$  i  $c$  oznaczoną symbolem  $\bar{S}_D$ . W metodzie tej również średnie wartości modułów różnic między oszacowaniami parametrów  $a^*$ ,  $b^*$  i  $c^*$  a ich wartościami teoretycznymi  $a$ ,  $b$  i  $c$  są wyraźnie bliższe zeru.

### PRZYKŁAD LICZBOWY

Na podstawie danych pobranych ze strony GUS<sup>1</sup> dotyczących stopy bezrobocia (w %) dla 380 powiatów (na dzień 31.01.2018 r.), przygotowano siatkę prawdopodobieństwa za pomocą nowej metody III (wykr. 8). Funkcja celu  $M_3$  (22) dla  $n = 380$  przyjmuje wartość najmniejszą dla wartości parametrów  $a^* = 4,719$ ;  $b^* = 1,288$ ;  $c^* = 2,408$  URG.

<sup>1</sup> <http://stat.gov.pl/obszary-tematyczne/rynek-pracy/bezrobocie-rejestrowane/liczba-bezrobotnych-zarejestrowanych-oraz-stopa-bezrobocia-wedlug-makroregionu-regionu-i-podregionu-stan-w-koncu-stycznia-2018-r-,2,66.html>.

**WYKR. 8. SIATKA PRAWDOPODOBIENSTWA URG UZYSKANA METODĄ III**

Źródło: jak przy wyk. 1.

Wykr. 8 przedstawia siatkę URG uzyskaną metodą III złożoną z prostej nachylonej do osi OX pod kątem  $45^\circ$  oraz z punktów, które dążą do tej prostej.

**Podsumowanie**

Mając do dyspozycji komputer można urealnić to, co kiedyś wydawało się niemożliwe. Poruszając się sprawnie po metodach numerycznych i dysponując wiedzą przedstawioną powyżej można zbudować siatkę prawdopodobieństwa URG zarówno nową metodą, jak i metodami klasycznymi.

Badania symulacyjne pokazały, że najlepszą efektywnością estymacji charakteryzuje się metoda III tworzenia siatki prawdopodobieństwa wykorzystująca wartości modalne funkcji gęstości statystyk pozycyjnych. Świadczą o tym najmniejsza średnia odległość oszacowań parametrów URG od ich wartości teoretycznych oraz średnie wartości modułów różnic między oszacowaniami parametrów  $a^*$ ,  $b^*$  i  $c^*$  URG a ich wartościami teoretycznymi  $a$ ,  $b$  i  $c$ . Zatem cel artykułu został osiągnięty.

## LITERATURA

- Chambers, J. M. (1983). *Graphical methods for data analysis*. Belmont: Wadsworth International Group.
- Cleveland, W. S. (1985). *The elements of graphing data*. Monterey: Wadsworth Advanced Books and Software.
- David, H. A. (1970). *Order statistics*. New York: Wiley.
- Drapella, A., Kosznik, S., Perzyńska, G. (1999). Nowy sposób określania położenia punktów doświadczalnych na siatce rozkładu Weibulla. *Śląskie Prace Matematyczno-Fizyczne*, (12a), 155—162.
- Kosznik, S. (1999). The modified Makeham lifetime model an overview. *Śląskie Prace Matematyczno-Fizyczne*, (12a), 177—194.
- Kroese, D. P., Taimre, T., dan Botev, Z. I. (2011). *Handbook of Monte Carlo Methods*. Hoboken: John Wiley & Sons.
- Stacy, E. W. (1962). A generalization of the gamma distribution. *The Annals of Mathematical Statistics*, 33(3), 1187—1192.
- Stacy, E. W., Mihram, G. A. (1965). Parameter estimation for a generalized gamma distribution. *Technometrics*, 7(3), 349—357.
- Sulewski, P. (2003). Nowa metoda wyznaczania położenia punktów doświadczalnych na siatce rozkładu normalnego. *Wiadomości Statystyczne*, 11, 1—9.
- Sulewski, P. (2008). *Uogólniony rozkład gamma w praktyce statystyka*. Śląsk: Wydawnictwo Naukowe Akademii Pomorskiej.
- Sulewski, P. (2014). Wykorzystanie uogólnionego rozkładu gamma do generowania tablicy dwudzielczej. *Śląski Przegląd Statystyczny*, 12(8), 339—347.
- Thode, H. C. (2002). *Testing for normality*. New York: Marcel Dekker.
- Wilk, M. B., Gnanadesikan, R. (1968). Probability Plotting Methods for the Analysis of Data. *Biometrika*, 55(1), 1—17.
- Zieliński, R. (1979). *Generatory liczb losowych*. Warszawa: WN-T.

**Summary.** *The aim of the paper is to propose a new method of creating a Q-Q plot using the density function of order statistics as well as to compare it with the classical methods. The most popular distributions for statisticians are those flexible ones which have easily estimated parameters. In the pre-computer era Quantile-Quantile plot (Q-Q plot) can be constructed only for distributions of reversible cumulative distribution functions (CDF) such as the exponential distribution and the Weibull distribution. The CDF of generalised gamma distribution (GGD) is not only analytically irreversible, but also has no analytical form. However, at present, owing to advanced computer technology, this problem can be solved. The CDF of GGD can be inverted by using different computing environment, i.e. Microsoft Excel, Mathcad, R language.*

**Keywords:** Q-Q plot, generalised gamma distribution, cumulative distribution function, Monte Carlo method.