

Sprawozdanie 14.

Generowanie ciągu liczb pseudolosowych o rozkładzie jednorodnym i trójkątnym

Mirosław Kołodziej

15.06.2021

1. Wstęp teoretyczny

1.1 Generatory liniowe

Generatory liniowe to generatory tworzące ciąg liczb według poniższego schematu:

$$X_{n+1} = (a_1 X_n + a_2 X_{n-1} + \dots + a_k X_{n-k+1} + c) \bmod m,$$

gdzie $a_1, a_2, \dots, a_k, c, m$ są parametrami generatora – to ustalone liczby.

Generatory liniowe dzielimy na dwa rodzaje:

- generator multiplikatywny dla $c = 0$,
- generator mieszany dla $c \neq 0$.

Maksymalny okres generatora liniowego to $m - 1$.

1.2 Rozkład trójkątny

Funkcja gęstości prawdopodobieństwa oznaczana jest w następujący sposób:

$$f(x; \mu, \Delta) = -\frac{|x - \mu|}{\Delta^2} + \frac{1}{\Delta},$$

gdzie μ to środek rozkładu, a Δ to jego szerokość. Dystrybucja tego rozkładu jest następująca:

$$F(a) = P(x < a) = \int_{\mu-\Delta}^a f(x; \mu, \Delta) dx = \begin{cases} -\frac{1}{\Delta^2} \left(-\frac{x^2}{2} + \mu x \right) + \frac{x}{\Delta}, & x \leq \mu \\ -\frac{1}{\Delta^2} \left(\frac{x^2}{2} - \mu x + \mu^2 \right) + \frac{x}{\Delta}, & x > \mu \end{cases}.$$

Dla $\xi_1 \in U(0,1)$ i $\xi_2 \in U(0,1)$ zmienne o rozkładzie trójkątnym oraz o parametrach μ i Δ generujemy za pomocą formuły:

$$x = \mu + (\xi_1 + \xi_2 - 1) \cdot \Delta$$

2. Problem

Celem naszych zajęć laboratoryjnych było wygenerowanie ciągu liczb pseudolosowych o rozkładzie jednorodnym i trójkątnym.

Zaczęliśmy od rozkładu jednorodnego. Startowaliśmy od $x_0 = 10$ i generowaliśmy $n = 10^4$ liczb pseudolosowych za pomocą generatora mieszanego:

$$x_{n+1} = (ax_n + c) \bmod n$$

o następujących parametrach dla każdej wersji:

- a) $a = 123, c = 1, m = 2^{15}$,
- b) $a = 69069, c = 1, m = 2^{32}$.

Dla obu podpunktów przygotowaliśmy rysunki $X_{i+1} = f(X_i)$ ($X_i = \frac{x_i}{m+1.0}$ z warunku normalizacji rozkładu $U(0,1)$). Korzystaliśmy przy tym z funkcji $\text{gen_1}()$.

Następnie przeszliśmy do rozkładu trójkątnego. Wygenerowaliśmy za jego pomocą 10^3 liczb, zaś parametry ustawiliśmy na $\mu = 4$ i $\Delta = 3$. Podzieliliśmy przedział $[\mu - \Delta, \mu + \Delta]$ na $K = 10$ podprzedziałów i zliczyliśmy ile liczb wpada do każdego z nich. Przeprowadziliśmy również test χ^2 , czyli określiliśmy wartość statystyki testowej:

$$\chi^2 = \sum_{i=1}^K \frac{(n_i - n \cdot p_i)^2}{n \cdot p_i},$$

gdzie n_i to ilość liczb znajdujących się w podprzedziale o indeksie i , p_i to prawdopodobieństwo teoretyczne ($p_i = F(x_{i,max}) - F(x_{i,min})$), że zmienna losowa X znajdzie się w i – tym przedziale.

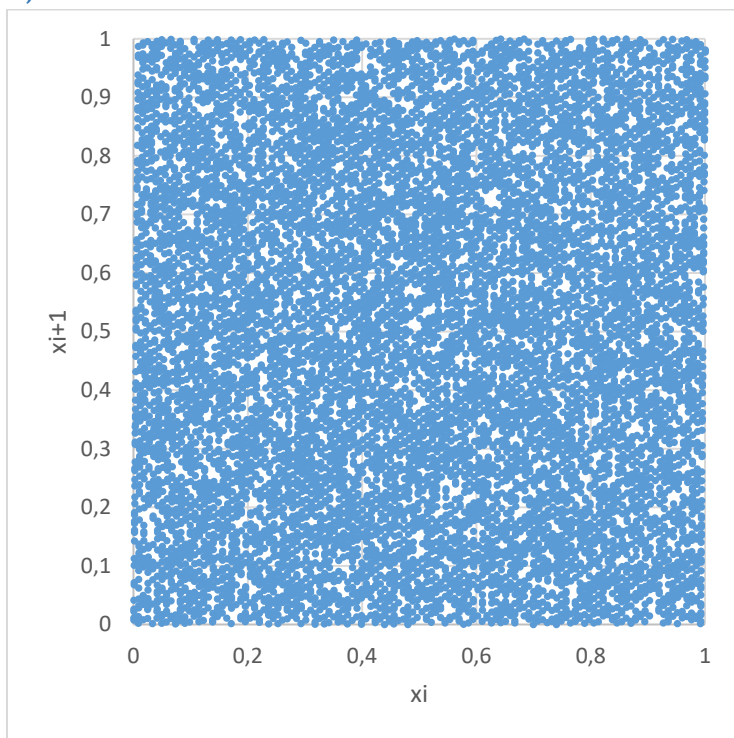
Na koniec testowaliśmy hipotezę H_0 , czyli czy wygenerowany rozkład jest rozkładem $T(\mu, \Delta)$ wobec H_1 , że nie jest to prawda. Z pomocą tabel statystycznych sprawdzaliśmy, czy Hipoteza jest prawdziwa na poziomie istotności $\alpha = 0.05$. W tym celu zdefiniowaliśmy obszar krytyczny testu:

$$\Phi = \{X: \chi^2(X) > \varepsilon\},$$

gdzie $X = \{x_1, x_2, \dots, x_n\}$ jest ciągiem liczb pseudolosowych, $\chi^2(X)$ jest wartością statystyki dla danego ciągu X , zaś ε to poziom krytyczny danego rozkładu dla określonej liczby stopni swobody i założonego poziomu istotności. Liczbę stopni swobody określaliśmy jako $\nu = K - r - 1$, gdzie K jest liczbą podprzedziałów, zaś $r = 2$ jest liczbą parametrów testowanego rozkładu (μ i Δ). Jeśli $\chi^2 < \varepsilon$ to stwierdzaliśmy, że dla zadanego poziomu istotności nie ma podstaw do odrzucenia hipotezy H_0 .

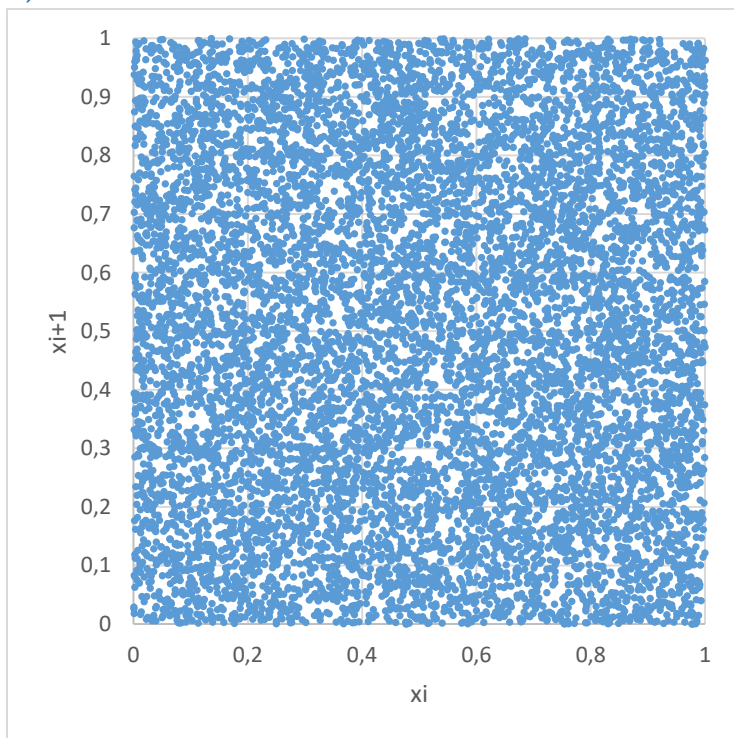
3. Wyniki

3.1 Zależność X_{i+1} z warunku normalizacji do rozkładu $U(0,1)$ dla $a = 123, c = 1, m = 2^{15}$



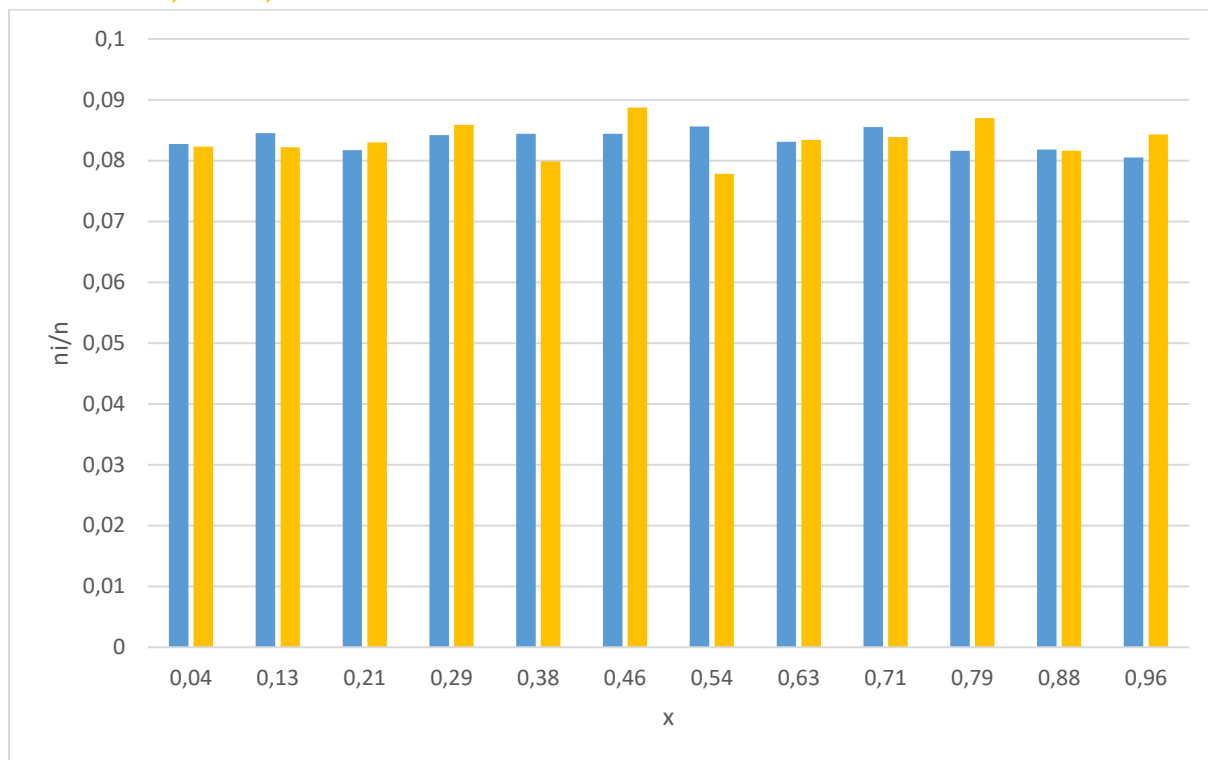
$\mu = 0.498215$ (średnia), $\sigma = 0.287120$ (odchylenie)

3.2 Zależność X_{i+1} z warunku normalizacji do rozkładu $U(0,1)$ dla $a = 69069, c = 1, m = 2^{32}$

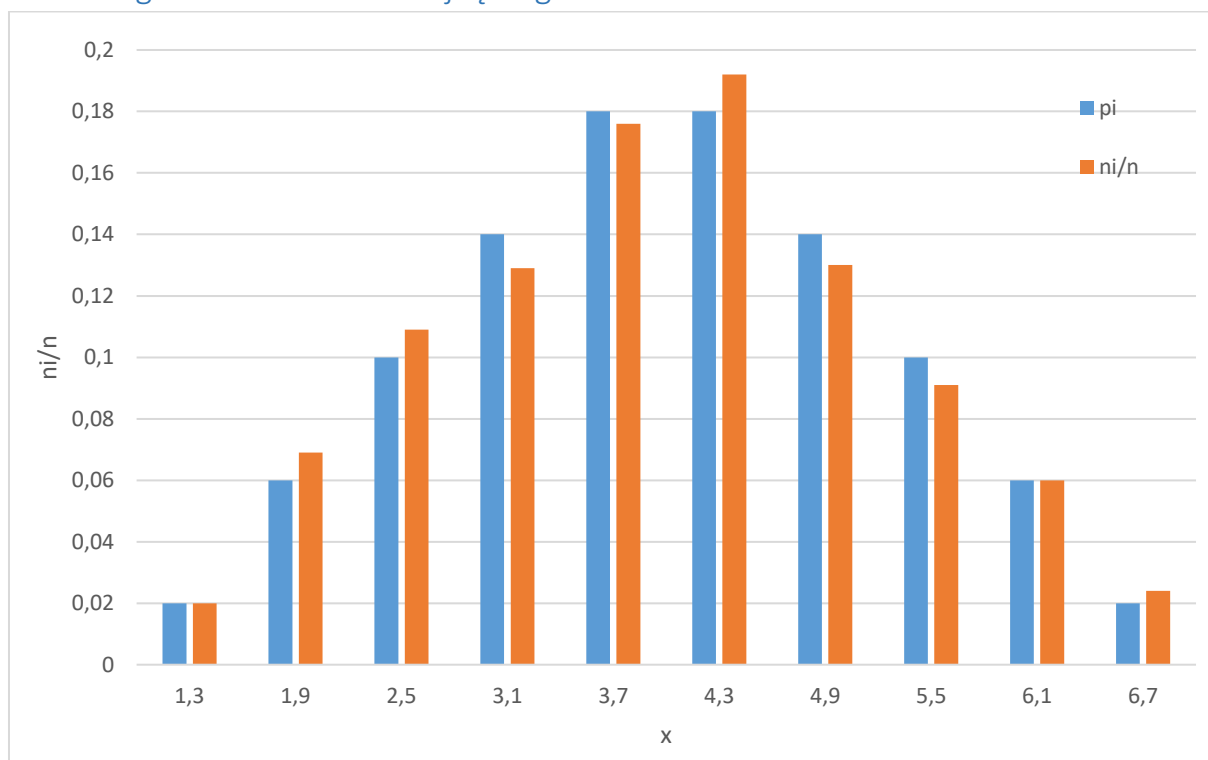


$\mu = 0.501622$ (średnia), $\sigma = 0.288302$ (odchylenie)

3.3 Histogram dla rozkładu jednorodnego dla $a = 123, c = 1, m = 2^{15}$ oraz dla $a = 69069, c = 1, m = 2^{32}$



3.4 Histogram dla rozkładu trójkątnego



3.5 Tabela przedstawiająca obliczoną wartość χ^2 oraz wartość tablicową ε

| χ^2 | ε |
|----------|---------------|
| 6.237460 | 14.07 |

4. Wnioski

Wygenerowaliśmy ciągi liczb pseudolosowych o rozkładzie jednorodnym i trójkątnym. Dla rozkładu jednorodnego możemy stwierdzić, że dla parametrów o wyższych wartościach zwiększona została losowość, jednak, co jest zauważalne na histogramie, w tym przypadku tworzą się większe odchylenia.

Dla rozkładu trójkątnego zauważmy również, że wartości nie do końca pokrywają się z teoretycznymi. Możemy dodatkowo stwierdzić, że hipoteza H_0 nie ma podstaw do odrzucenia.