Sprawozdanie XIII

Generowanie liczb pseudolosowych o rozkładzie jednorodnym i trójkątnym

Adam Łaba

10 czerwca 2021

1. Wstęp teoretyczny

1.1. Rozkład jednorodny

Rozkład jednorodny to rozkład prawdopodobieństwa, dla którego funkcja rozkładu ma stałą wartość. W praktyce można uzyskać zmienne rozkładu jednorodnego wykorzystując generatory liczb pseudolosowych. W tym ćwiczeniu zostały wykorzystane generatory mieszane, które następnie normalizowano do rozkładu $U(0, 1)$ w ten sposób dając rozkład jednorodny. Generator mieszany:

$$X_{n+1} = (aX_n + c)\, mod\, m$$

Wzór 1.

Normalizacja:

$$x_i = \frac{X_i}{m+1.0}$$

Wzór 2.

Dodatkowo, średnia N elementów:

$$\mu = \frac{1}{N}\sum_{i=0}^{N-1} x_i$$

Wzór 3.

Odchylenie standardowe:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1} (x_i - \mu)^2}$$

Wzór 4.

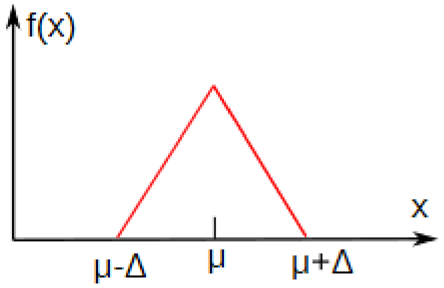
1.2. Rozkład trójkątny

Jest definiowany przez funkcję gęstości prawdopodobieństwa wyglądającą następująco:

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

Wzór 5.

Gdzie μ to środek rozkładu, a Δ to jego szerokość. Funkcja na rysunku prezentuje się:



Rys1. Rozkład trójkątny

Dystrybuanta dla rozkładu trójkątnego:

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

Wzór 6.

1.3. Test chi-kwadrat

Test służy do badania hipotezy, że pewna zmienna X ma rozkład prawdopodobieństwa o dystrybuancie F. Wartość statystyki testowej:

$$\chi^2 = \sum_{j=0}^{K-1} \frac{(n_j - Np_j)}{Np_j}$$

Wzór 7.

Gdzie p_j to teoretyczne prawdopodobieństwo wylosowania liczby z j-tego przedziału:

$$p_j = F(x_{j,min}) - F(x_{j,min})$$

Wzór 8.

A $F(x)$ to dystrybuanta rozkładu.

2. Zadanie do wykonania

2.1. Rozkład jednorodny

Pierwsza część zadania polegała na generowaniu pseudolosowych liczb rozkładu jednorodnego poprzez wykorzystanie generatora mieszanego a następnie normalizację wartości. W każdym przypadku ziarnem generatora było $x_0 = 10$, generowaliśmy $n = 10^4$ liczb. Całość powtórzono dla dwóch generatorów o parametrach:

a) $a = 123$, $c = 1$, $m = 2^{15}$

b) $a = 69069$, $c = 1$, $m = 2^{32}$

Na podstawie otrzymanych wartości przedstawiono na wykresie zależność między elementami x_i a x_{i+1}. Dodatkowo obliczono średnią i odchylenie standardowe dla liczb wygenerowanych w obu przypadkach celem porównania ich z wartościami teoretycznymi oraz z wartościami obliczonymi dla funkcji rand() z języka C. Następnie podzielono przedział (0, 1) na 12 podprzedziałów i zliczano ilość wystąpień w danym podprzedziale celem zobrazowania działania generatorów na histogramie.

2.2. Rozkład trójkątny

Kolejnym zadaniem było wygenerowania $n = 10^3$ liczb podlegających rozkładowi trójkątnemu o parametrach $\mu = 4$, $\Delta = 3$. Wartości generowano w oparciu o wzór:

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

Wzór 9.

Gdzie ξ_1, ξ_2 - liczby losowe z przedziału (0,1) uzyskane przy wykorzystaniu generatora analogicznego jak z ***2.1.b)*** (zaczynając od ziarna $x_0 = 10$). W tym przypadku również podzielono przedział na podprzedziały (tym razem 10) i zliczano ilość wystąpień w celu przedstawienia wyników na histogramie. Dodatkowo w oparciu o te wyniki przeprowadzono test chi-kwadrat.

Na podstawie wyników testu przeprowadzono następnie testowanie poniższej hipotezy:
H_0: *Wygenerowany rozkład jest rozkładem T(μ, Δ) wobec* H_1 *że nie jest to prawdą.*
Przyjęty poziom istotności $\alpha = 0.05$. Obszar krytyczny testu:

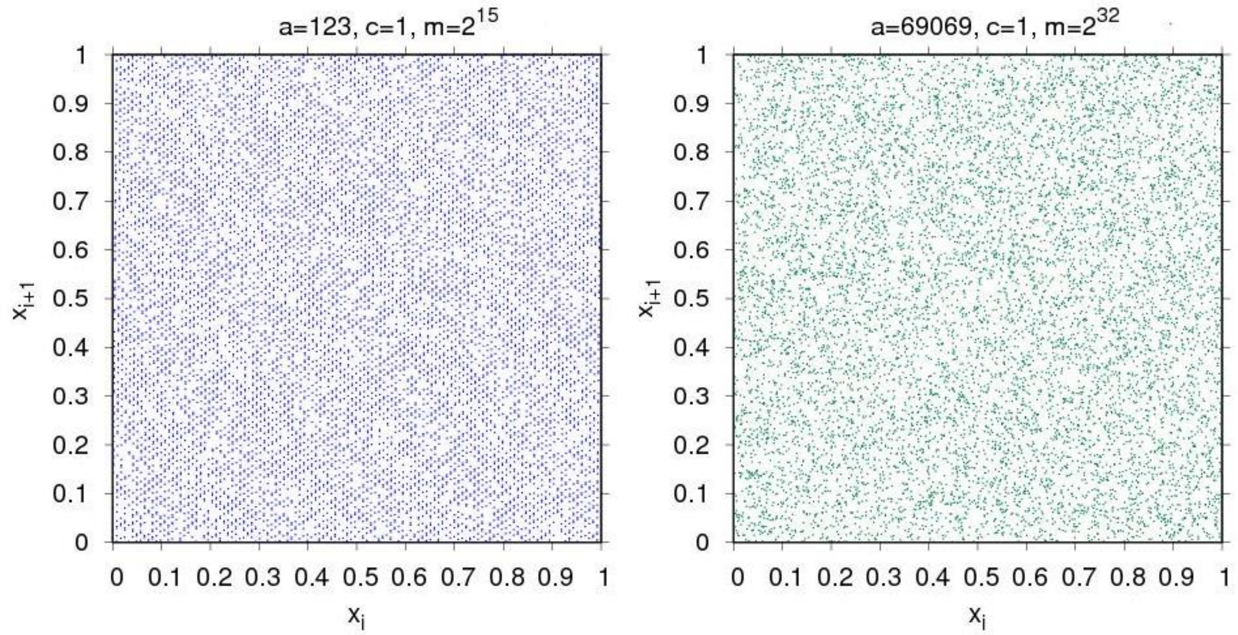
$$\Phi = \left\{ X : \chi^2(X) > \epsilon \right\}$$

Wzór 10.

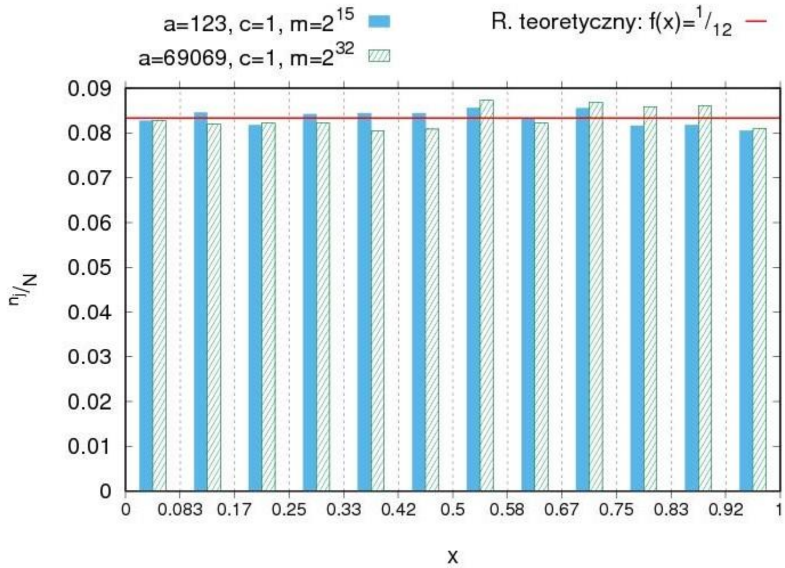
Gdzie: X - ciąg liczb pseudolosowych, $\varepsilon = 14.06$ - poziom krytyczny (wartość odczytana z tablic).

3. Wyniki

3.1. Zależność między x_i a x_{i+1} dla generatorów rozkładu jednorodnego



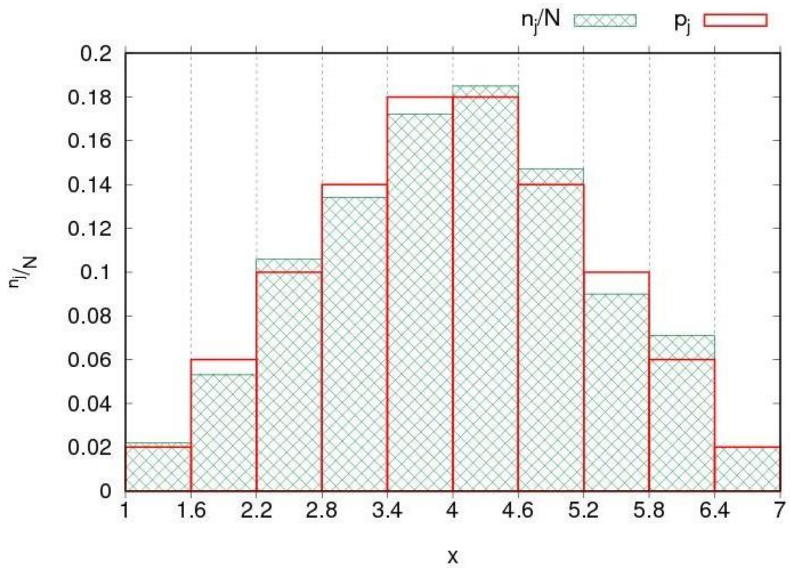
3.2. Histogramy generatorów rozkładu jednorodnego



3.3. Wartości średnich i odchyleń standardowych dla wygenerowanych liczb

wartość	teoretyczna	generator a)	generator b)	rand()
μ	0.5	0.498266	0.503806	0.497132
σ	$\sqrt{\frac{1}{12}}$	0.287120	0.288070	0.288266

3.4. Histogram rozkładu trójkątnego



3.5. Test chi-kwadrat i testowanie hipotezy

Uzyskany wynik testu:

$$\chi^2 = 5.494921$$

W związku z czym:

$$\chi^2 < \epsilon \qquad (5.494921 < 14.06)$$

Więc hipoteza nie została odrzucona.

4. Wnioski

Generatory mieszane już dla stosunkowo małych wartości *a* (123) i *m* (2^{15}) generują liczby pseudolosowe które po unormowaniu dają wartości μ i σ bliskie teoretycznym, lecz zauważalna jest zależność między kolejnymi elementami. Wraz ze wzrostem wartości *a* i *m* zależność między kolejnymi elementami jest coraz mniej widoczna. Wyniki μ i σ dla obu generatorów są porównywalne z wynikami funkcji rand().

Histogramy dla dwóch pierwszych generatorów obrazują bliskość generowanego rozkładu do rozkładu jednorodnego.

Generowany rozkład trójkątny również był bliski teorii – jest to widoczne na histogramie, reprezentującym charakterystyczny kształt. Uzyskany wynik testu chi-kwadrat wykorzystany przy testowaniu hipotezy z przyjętym poziomem istotności nie pozwolił na odrzucenie hipotezy.