SPRAWOZDANIE METODY NUMERYCZNE

LABORATORIUM NR 14 – GENEROWANIE CIĄGU LICZB PSEUDOLOSOWYCH O ROZKŁADZIE JEDNORODNYM I TRÓJKĄTNYM.

Aleksandra Krzemińska,

Informatyka Stosowana II rok WFiIS, I stopień,

rok 2021

Wstęp teoretyczny

Generator liczb pseudolosowych to algorytm, który na podstawie ‘ziarna’ (z ang. seed), czyli małej ilości danych wejściowych generuje ciąg bitów, który jest nieodróżnialny od ciągu wygenerowanego z prawdziwie losowego źródła.

Generatory liniowe towrzą ciąg liczb według shematu:

$$X_{n+1} = (a_1 X_n + a_2 X_{n-1} + \cdots + a_k X_{n-k+1} + c) \, mod \, m \quad , \tag{1}$$

gdzie a1, a2, … ak, c są ustalonymi liczbami. Generator ten ma dwie odmiany – multiplikatywny, gdy c jest równe 0, lub mieszany - gdy c jest różne od 0.

Funckja gęstości prawdopodobieństwa dla rozkładu trójkątnego prezentuje się jak poniżej:

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

a dystrybuanta tego rozkładu jest następująca:

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

Na tej podstawie, można otrzymać formułę dzięki której uda się wygenerować liczby właśnie o rozkładzie trójkątnym. Można ją opisać jako:

$$x = \mu + (\xi_1 + \xi_2 - 1) \cdot \Delta \quad . \tag{3}$$

Zadanie do wykonania

Zadaniem laboratoryjnym było:

1) Wygenerować 10^4 liczb przy pomocy generatora mieszanego,
2) Wygenerować 10^3 liczb o rozkładzie trójkątnym, następnie podzielić przedział losowania na 10 poprzedziałów i sprawdzić ile liczb wpada do poszczególnych podprzedziałów.
3) Dla rozkładu trójkątnego określić wartość statystyki testowej:

$$\chi^2 = \sum_{i=1}^{K} \frac{(n_i - n \cdot p_i)^2}{n \cdot p_i} \quad , \tag{4}$$

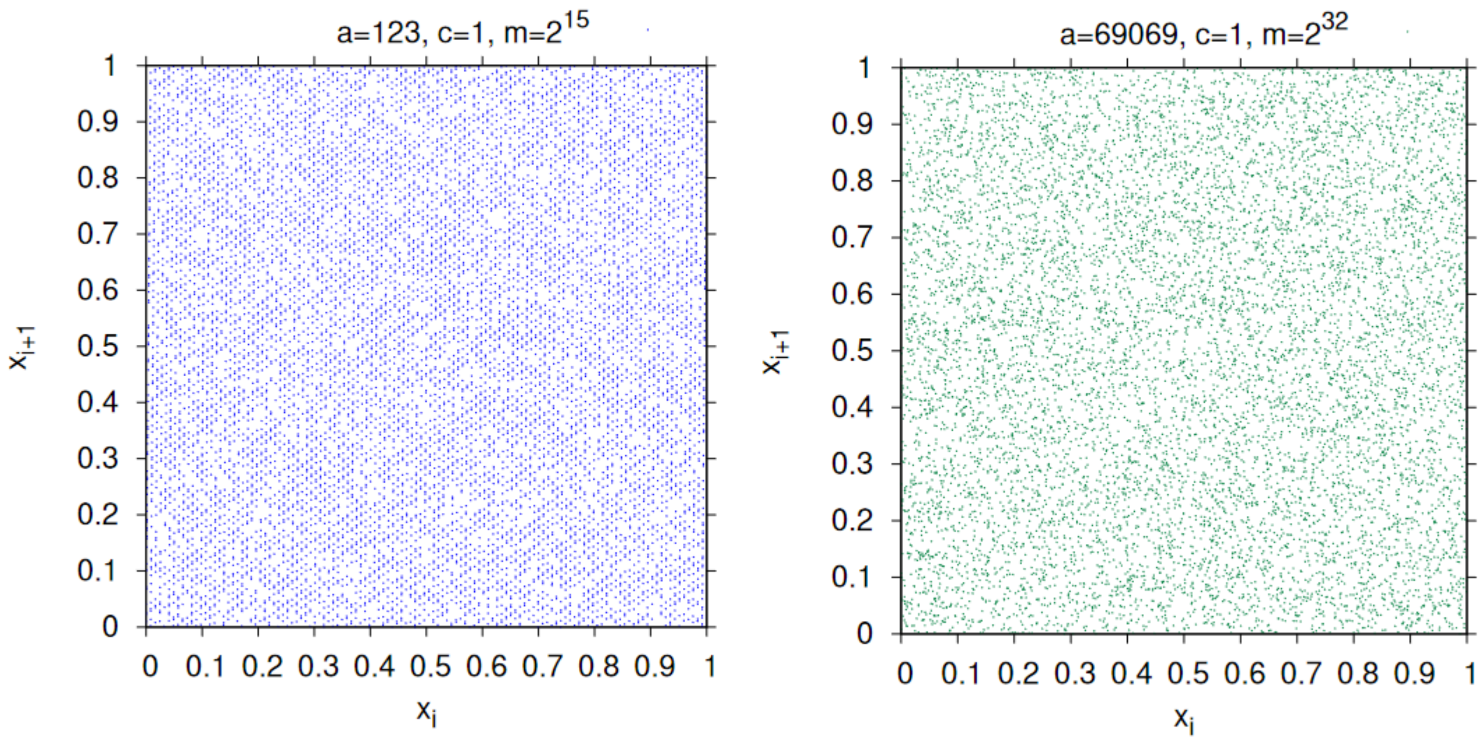
a następnie przetestować hipotezę dla zadanej istotności.

METODA

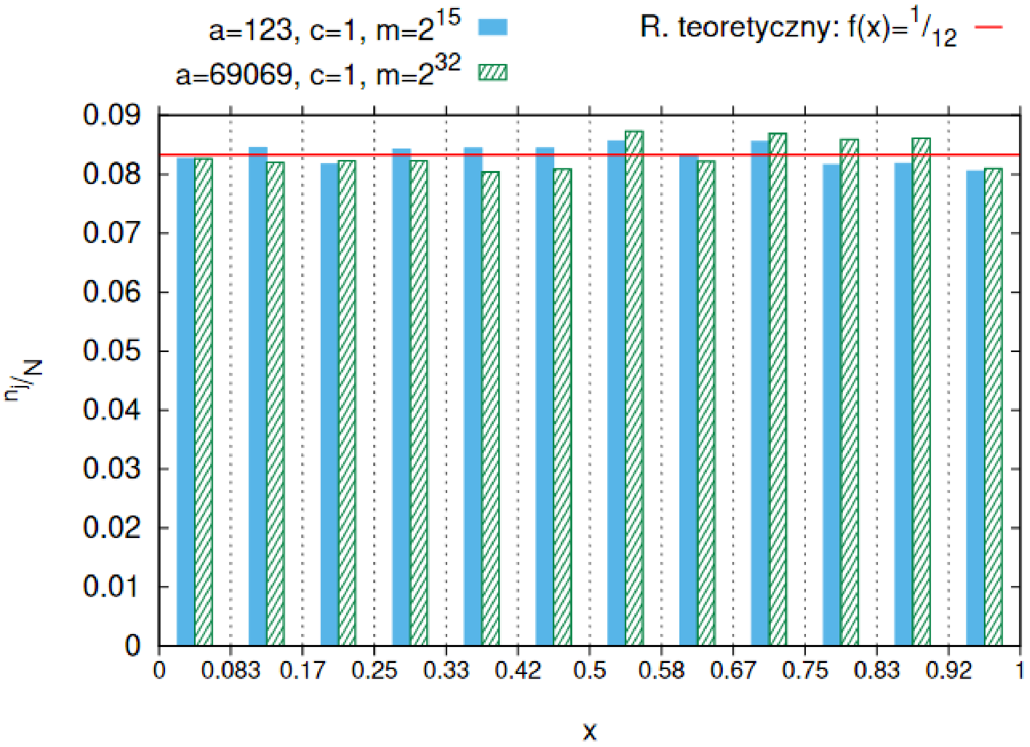
Dla podpunktu pierwszego zostały przyjęte dane w dwóch wersjach: 1) a = 123, c =1, m = 215 oraz 2) a = 69069, c = 1, m = 232. W podpunkcie drugim przyjęte zostały zmienne μ = 4 oraz Δ = 3. Wykorzystany został *wzór 3,* w którym wartości $\xi_{1,2}$ zostały wygenerowane za pomocą generatora z wersji *nr 2)* podpunktu 1. Następnie postępowanie było zgodne z zadaną instrukcją wraz ze wzorami przedstawionymi we wstępie teoretycznym. Pod koniec została przyjęta hipoteza, iż wygenerowany przez algorytm (w punkcie 2.) rozkład jest trójkątny - poziom istotności dla badanej hipotezy został ustalony jako α = 0.05.

WYNIKI

Wyniki zostały zaprezentowane poniżej.



Rysunek 1: Porównanie zależności x(i+1) od x(i) dla generatorów z punktu 1. – po lewej *wersja 1)* po prawej *wersja 2).*



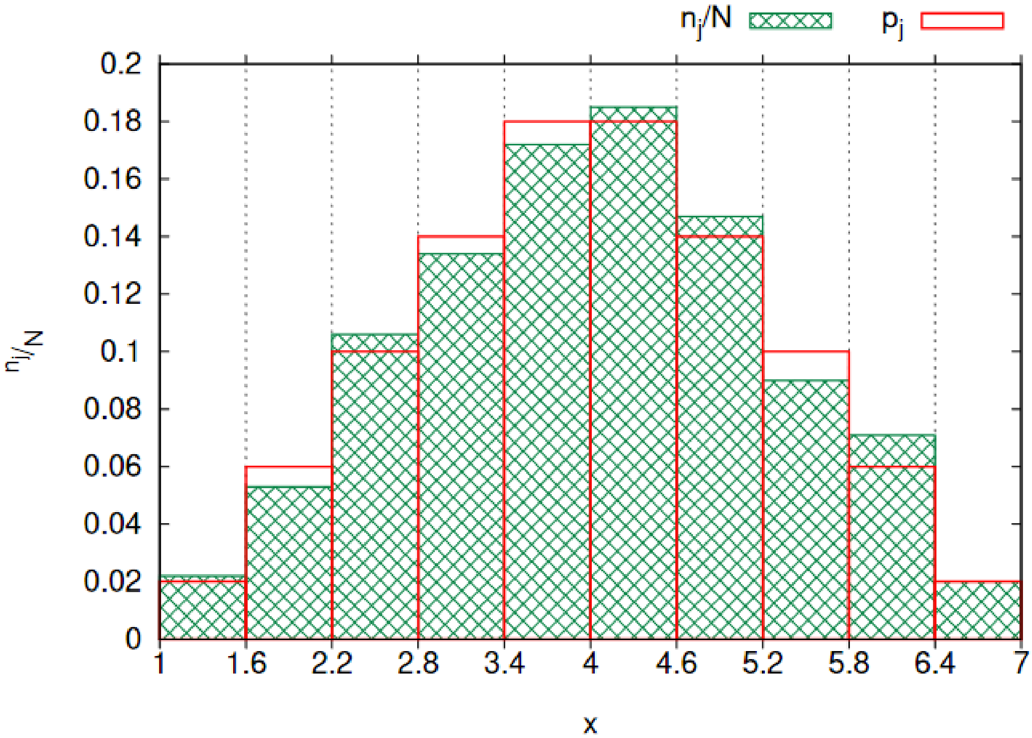
Rysunek 2: Histogram dla rozkładów pochodzących z obu generatorów mieszanych z punktu 1. – zielonym kolorem zaznaczona *wersja 1)*, niebieskim - *wersja 2)*.

W ramach weryfikacji wyników zostały także obliczone średnie oraz odchylenie standardowe dla obu generatorów pseudolosowych z podpunktu 1. Wyniki przedstawione zostały w poniższej tabeli i zestawione z wartościami teoretycznymi.

Tabela 1: Porównanie wyników średniej i odchylenia standardowego dla generatorów mieszanych.

Wersja	Średnia μ	Wartość teoretyczna μ	Odchylenie standardowe σ	Wartość teoretyczna σ
1 (a = 123, c = 1, m = 2^{15})	0.498266	0.500000	0.287120	0.288675
2 (a = 69069, c = 1, m = 2^{32})	0.503806	0.500000	0.288070	0.288675

Poniżej prezentują się wyniki dla generator o rozkładzie trójkątnym.



Rysunek 3: Histogram dla wygenerowanego rozkładu trójkątnego – przedstawia ilość wylosowanych liczb w j-tych przedziałach, a pj to teoretyczne prawdopodobieństwo.

Wyliczona ze *wzoru 4* statystyka testowa wyniosła 5.49492, zatem korzystając z podanych tablic statystycznych hipoteza nie została odrzucona na zadanym poziomie istotności ($\alpha = 0.05$).

Wnioski

Wyniki jednoznacznie mogą świadczyć o 'losowości' wyboru kolejnych generowanych liczb. Zaprezentowane jest to na *Rysunku 1*, gdzie widać całkowity brak uzależnienia od siebie kolejnych generowanych liczb.

Analizując *Rysunek 2* można wnioskować, że na generację kolejnych liczb, nie ma znaczącego wpływu dobór stałej a oraz m – widać, że zarówno dla generatora w *wersji 1* jak i w *wersji 2* są niewielkie odstępstwa od przewidywanego, w pełni losowego rozkładu, ale wyniki te nie różnią się znacząco.

Na podstawie *Tabeli 1*, można też potwierdzić efektywność zaprogramowanego w ramach laboratorium generatora – wartości średniej i odchylenia standardowego w obu przypadkach są bardzo zbliżone do wartości teoretycznych. Trzeba jednak tu zaznaczyć, że w tym przypadku *wersja 2* (z większym parametrem a i m), jest bardziej zbliżona do teoretycznego rozkładu normalnego, lecz jest to różnica na poziomie jednej tysięcznej.

Histogram przedstawiony na Rysunku 3 przedstawia niewielkie odstępstwa teoretycznego rozkładu trójkątnego od tego wygenerowanego przez algorytm. Analizując jednak dalej, okazuje się, że dla danej liczby stopni swobody (w bieżącym przypadku równym 2) uzyskana statystyka testowa mieści się w istotności 0.05 - uzyskana wartość statystyki jest mniejsza od wartości tabelarycznej (5.49492 < 5.991). Wynika z tego, że generator – zgodnie z założeniem – wygenerował zbiór liczb o rozkładzie trójkątnym.

Podsumowując, można potwierdzić, że zaprogramowane w ramach zajęć generatory liczb pseudolosowych zadziałały zgodnie z założeniem oraz w pełni efektywnie.