

Imię i nazwisko	Sylwester Macura
Kierunek	Informatyka Stosowana
Rok	3
Grupa	2
Temat	Generowanie ciągu liczb pseudolosowych o rozkładzie normalny metodą eliminacji.

1. Wstęp teoretyczny

Generatory liczb losowych możemy podzielić na dwa rodzaje fizyczne i komputerowe. Zaletą fizycznych generatorów liczb losowych jest niezależność i brak korelacji pomiędzy wygenerowanymi liczbami losowymi. Wadą natomiast jest brak powtarzalności serii. Z drugiej strony generatory komputerowe posiadają tę zaletę że są powtarzalne.

Do najprostszych generatorów należą generatory liczbowe. Tworzą one kolejne liczby losowe według wzoru.

$X_{n+1} = (a_1 X_n + a_2 X_{n-1} + \dots + a_k X_{n-k+1} + c) \bmod m$ Gdzie $a_1, a_2, \dots, a_k, c, m$ są ustalonymi liczbami. Generatory takie nazywamy kongurencyjne. Najprostszy generator liniowy ma dwie odmiany multiplikatywny gdy $c=0$ i mieszany gdy $c \neq 0$.

Okres generatora multiplikatywnego $T = \min(i : X_i = X_0, i > 0)$

Maksymalny okres $a^{\frac{m-1}{p}} \neq 1 \bmod m$

Maksymalny okres generatora uzyskamy gdy m jest liczbą pierwszą a jest czynniki pierwszym liczby $(m-1)$.

Wadą tych generatorów jest nierówne pokrycie przestrzeni. Generowane liczby układają się na hiperprzestrzeniach które są uzależnione od parametrów generatora.

Aby sprawdzić poprawność generatora o rozkładzie równomiernym w $(0,1)$ musimy wyznaczyć wartość oczekiwaną oraz wariancję ciągu liczb wyjściowych.

Wartość oczekiwana $\bar{\mu} = \frac{1}{N} \sum_{i=1}^N x_i$ powinna być równa $1/2$ oraz odchylenie standardowe

$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$ wynosić $1/12$. Dodatkowo współczynnik autokorelacji powinien być równy

0.

Aby uzyskać ciąg liczb losowych możemy posłużyć się metodą eliminacji, wykorzystujemy wtedy schemat.

1. Losujemy dwie zmienne o rozkładzie równomiernym $U_1 \in [a, b]$ $U_2 \in [0, d]$
2. Jeżeli $U_2 \leq F(U_1) \Rightarrow X = U_1$
3. Gdy powyższy warunek nie jest spełniony odrzucamy liczby
4. Wykonujemy czynności 1-3 aż do uzyskania odpowiedniej ilości liczb

Funkcję gęstości prawdopodobieństwa rozkładu normalnego definiujemy następująco

$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ natomiast dystrybucję wyznaczamy za pomocą wzoru

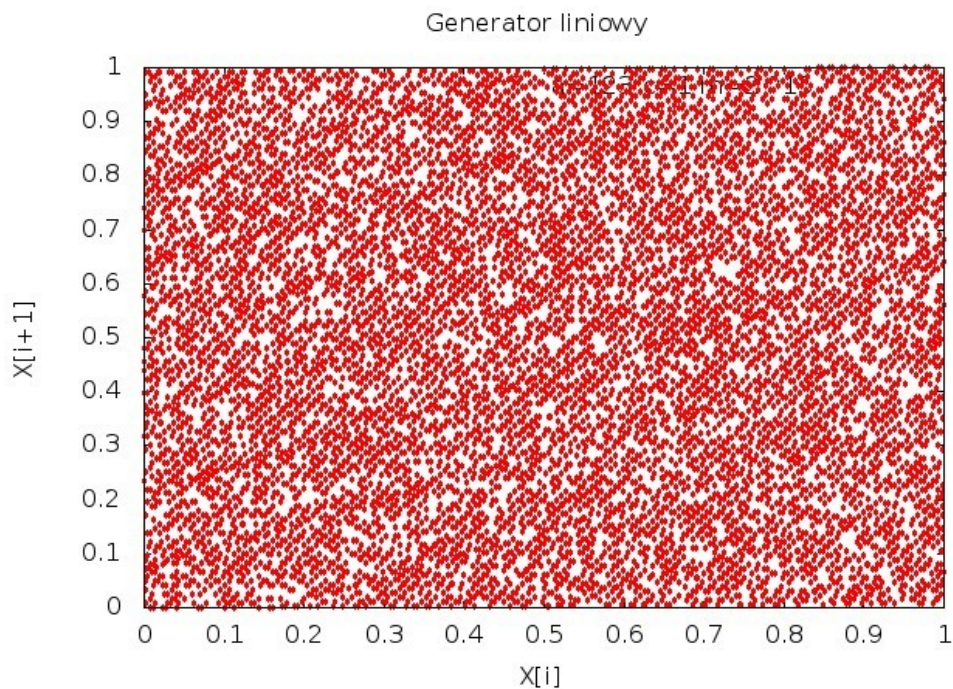
$F(x) = \frac{1 + \operatorname{erf}(x')}{2}$ gdzie $x' = \frac{x-\mu}{\sqrt{2}\sigma}$

2. Zadania do wykonania.

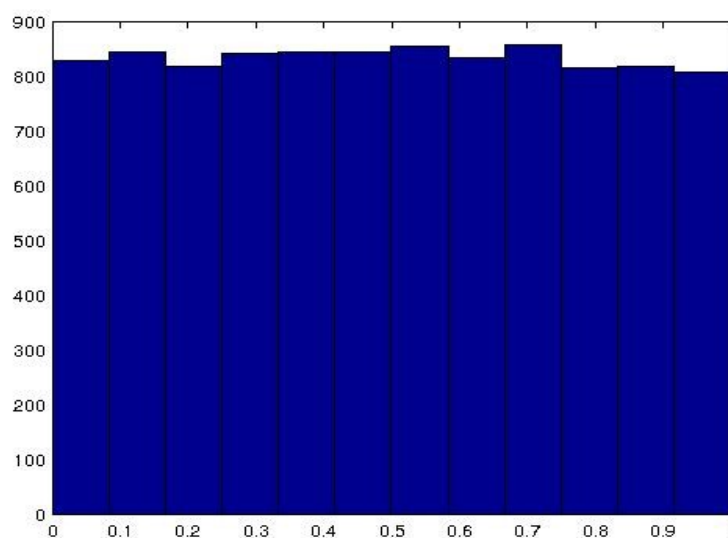
- Wygenerować ciąg liczb pseudolosowych za pomocą generatora liniowego o parametrach $a=123$ $c=1$ $m=2^{15}$ sporządzić rysunek zależności między zmiennymi losowymi oraz histogram
- Wygenerować ciąg liczb pseudolosowych za pomocą generatora liniowego o parametrach $a=69069$ $c=1$ $m=2^{32}$ sporządzić rysunek zależności między zmiennymi losowymi oraz histogram
- przetestować generator liczb pseudolosowych o rozkładzie normalnym

3. Wykonanie zdania

- Generator liniowy o parametrach $a=123$ $c=1$ $m=2^{15}$ $\mu=0.49$ $\sigma^2=0.08$

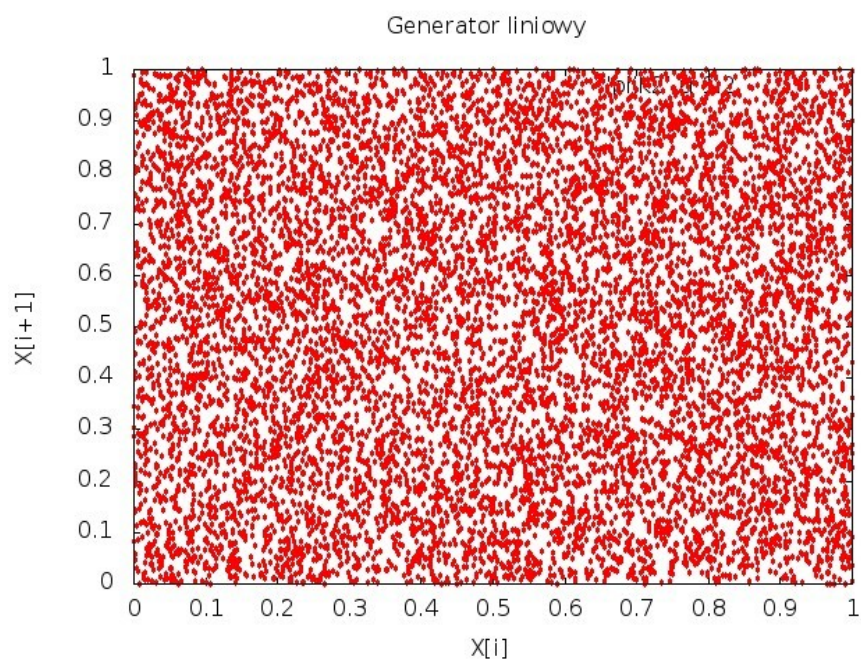


Ilustracja 1: Zależność między zmiennymi losowymi

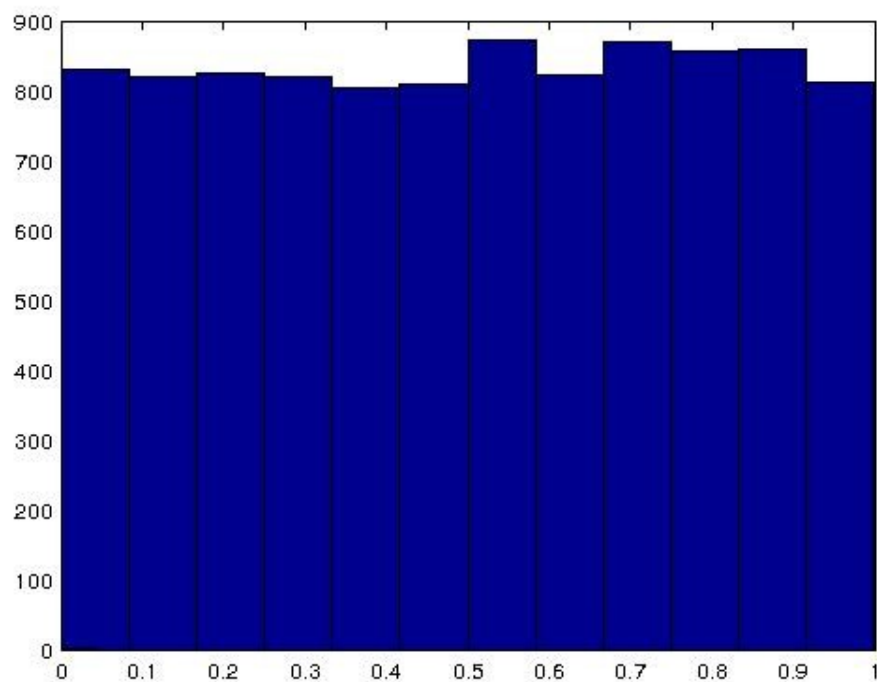


Ilustracja 2: Histogram

- Generator liniowy o parametrach $a=69069$ $c=1$ $m=2^{32}$ $\mu=0.51$ $\sigma^2=0.08$

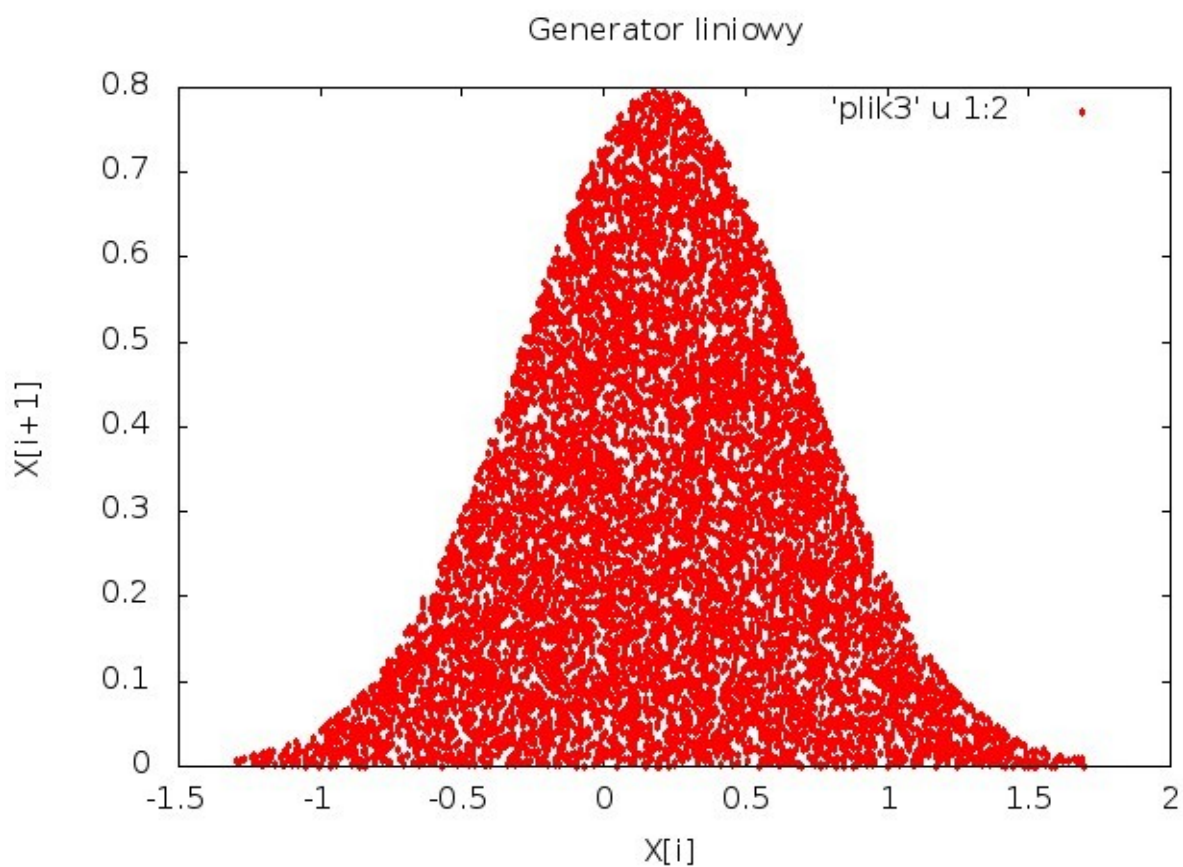


Ilustracja 3: Zależność między kolejnymi zmiennymi losowymi

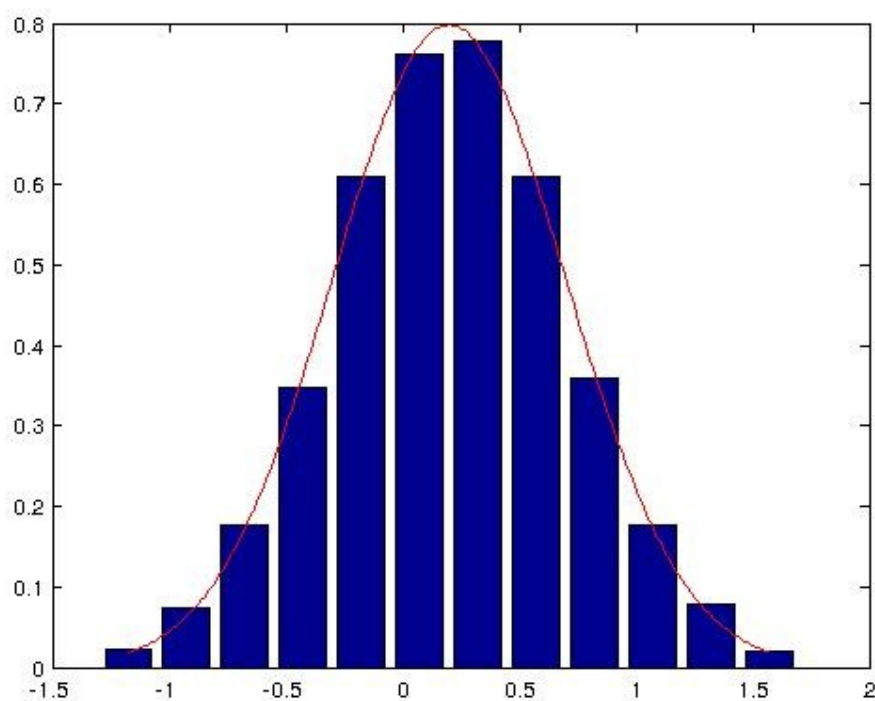


Ilustracja 4: Ilość punktów w każdym przedziale

- Generator o rozkładzie Gaussa o parametrach $a=69069$ $c=1$ $m=2^{32}$ $\mu=0.20$ $\sigma^2=0.24$



Ilustracja 5: Rozkład wygenerowanych punktów



Ilustracja 6: Stosunkowa ilość w poszczególnych pod przedziałach

4. Wnioski

Jak widzimy drugi generator liniowy ma lepsze parametry statystyczne od pierwszego. Dzieje się tak ponieważ posiada większy okres oraz używamy większej liczby pierwszej jako mnożnik. Pozwala to uzyskać w miarę nieskorelowane liczby. Jak widzimy wartości oczekiwane jak i wariancje uzyskane są bardzo bliskie wartościom teoretycznym.