Sprawozdanie

Identyfikacja i modelowanie statystyczne

Modelowanie i identyfikacja

Marcin Bober, 249426

Prowadzący:
Mgr inż. Maciej Filiński

22 marca 2022

Spis treści

1 Generator liczb pseudolosowych **2**
- 1.1 Opis . . . 2
- 1.2 Wpływ wartości początkowej X na własności generatora . . . 2
- 1.3 Wpływ parametru Z na własności generatora . . . 4
- 1.4 Okres generatora dla wybranych wartości Z . . . 6
- 1.5 Podobieństwo histogramu ciągu wygenerowaych liczb, a gęstość rozkładu jednostajnego . . . 6

2 Generator dany równaniem **7**
- 2.1 Opis . . . 7
- 2.2 Zależność od współczynnika m . . . 7
- 2.3 Zależność od współczynnika k . . . 8

3 Metoda odwracania dystrybuanty **10**
- 3.1 Opis . . . 10
- 3.2 Rozkład numer 1 . . . 10
- 3.3 Rozkład numer 2 . . . 11
- 3.4 Rozkład wykładniczy . . . 11
- 3.5 Rozkład Laplace'a . . . 12

4 Metoda odrzucania **13**
- 4.1 Rozkład numer 1 . . . 13
- 4.2 Rozkład numer 2 . . . 15
- 4.3 Półokrąg . . . 16
- 4.4 Rozkład Normalny . . . 18
- 4.5 Wnioski . . . 19

5 Estymacja **19**

6 Podsumowanie **19**

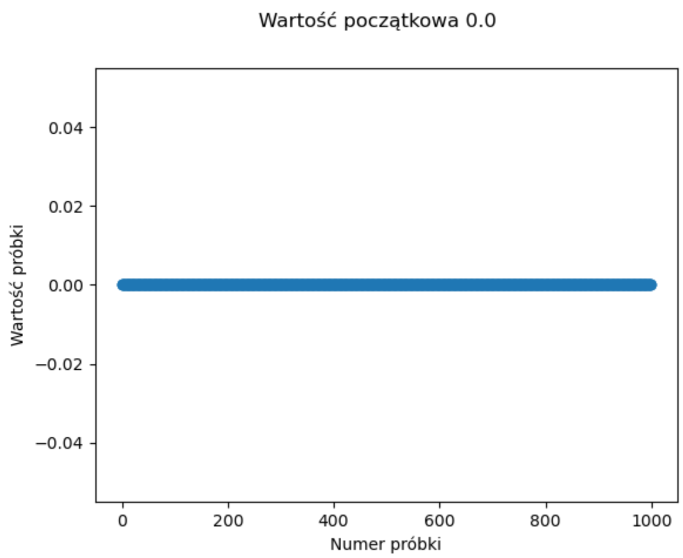
1 Generator liczb pseudolosowych

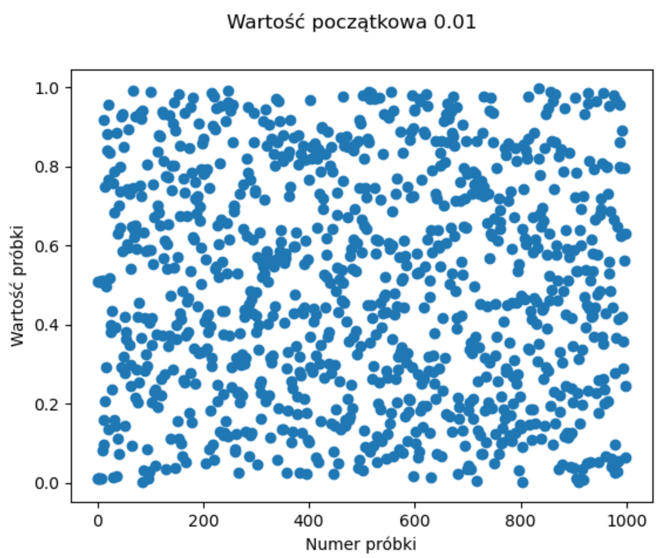
1.1 Opis

Zadanie polega na implementacji generatora liczb pseudolosowych z rozkładu jednostajnego oraz analizie wyników uzyskanych z jego udziałem. Generator oparty jest na przekształceniu piłokształtnym o równaniu $X_{n+1} = X_n \cdot z - [X_n \cdot z]$

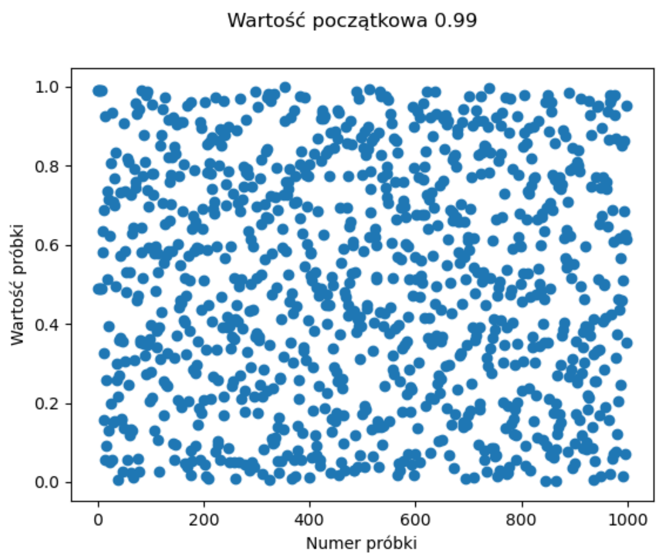
1.2 Wpływ wartości początkowej X na własności generatora

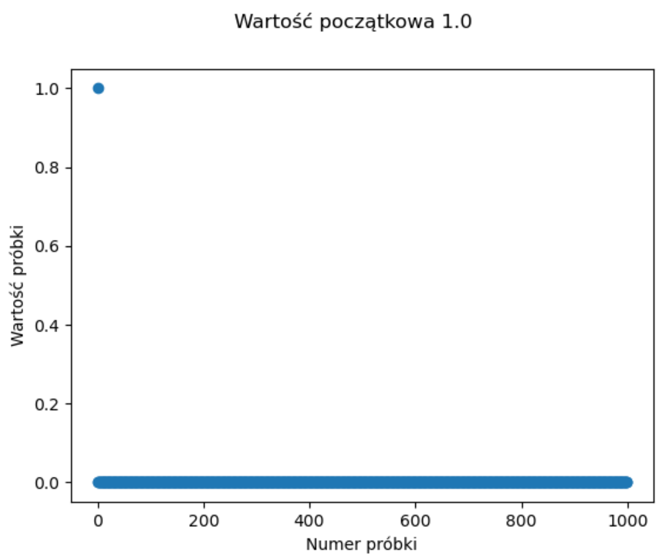
Wartość Z ustawiona została na wartość 51. Wykorzystano 1000 próbek.






Wartość początkowa 0.1
Wartość próbki
Numer próbki

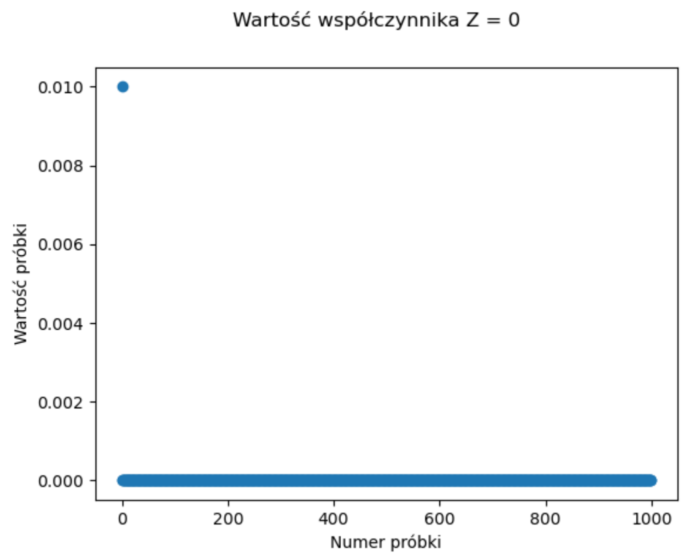

Wartość początkowa 0.99
Wartość próbki
Numer próbki

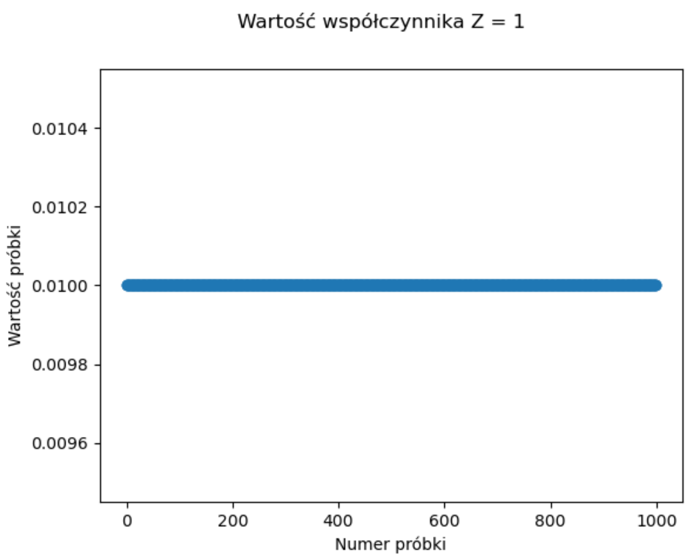

Wartość początkowa 1.0
Wartość próbki
Numer próbki

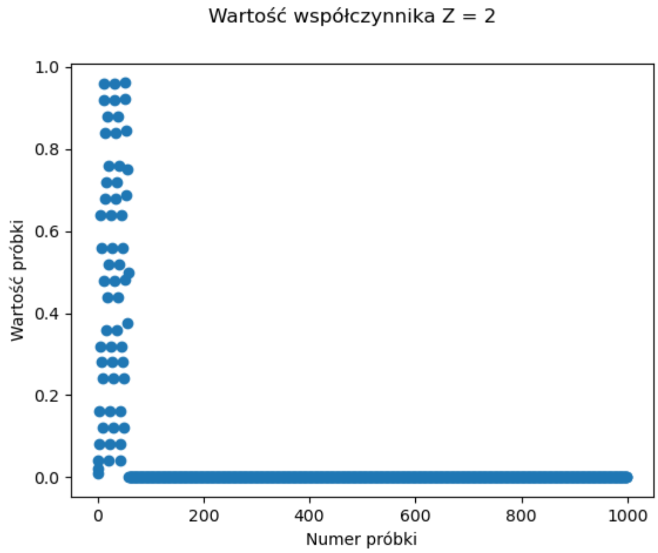
- Ustawienie wartości początkowej równej zero powoduje że wszystkie wygenerowane próbki są zerowe. (Patrz wykres 1.2) Dzieje się tak ponieważ algorytm opiera się o obliczenie iloczynu liczb, których jednym ze składników jest zero.
- Wybór liczby całkowitej spowoduje że pierwsza próbka jest równa tej wartości, a wszystkie kolejne są zerowe (Patrz wykres 1.2). Wynika to z faktu że obliczana jest reszta z dzielenia wartości przez jeden, która w taki wypadku zawsze równa jest zero.
- Zalecanym zakresem wyboru wartości początkowej jest przedział zawierający liczby większe od zera, z pominięciem liczb całkowitych.

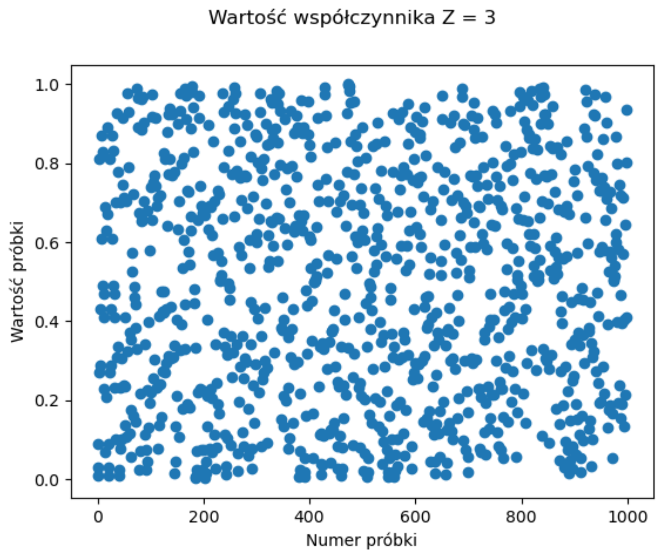
1.3 Wpływ parametru Z na własności generatora

Wartość X_0 ustawiona została na wartość 0,01. Wykorzystano 1000 próbek.









- Dla zerowego współczynnika Z pierwsza próbka uzyskuje wartość początkowa, a kolejne są zerami. Wynika to z mnożenia tych wyników przez współczynnik Z czyli zero.

- Gdy wartość Z jest równa jedności, wszystkie otrzymane wyniki są identyczne z wartością startową.

- W przypadku wykorzystania liczb parzystych, uzyskiwane wyniki szybko trafiają na wartość zero, która powoduje zatrzymanie generowania kolejnych wartości losowych.

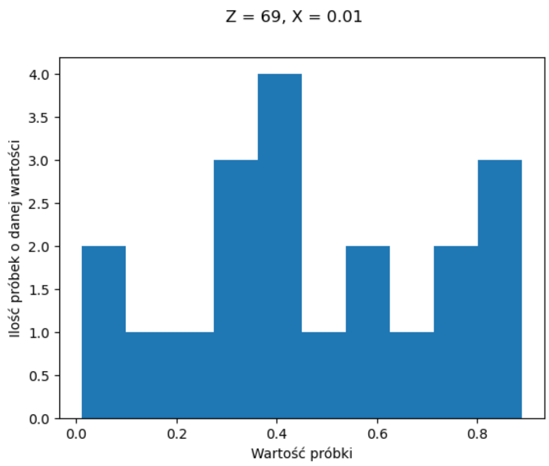
- Najlepsze wyniki otrzymywane są dla współczynnika Z będącego dużą liczbą pierwszą.

X_0	Z	okres generatora
0,1	1	1
0,1	2	4
0,1	3	4
0,1	4	2
0,1	5	1
0,1	6	1
0,1	7	4
0,1	8	4

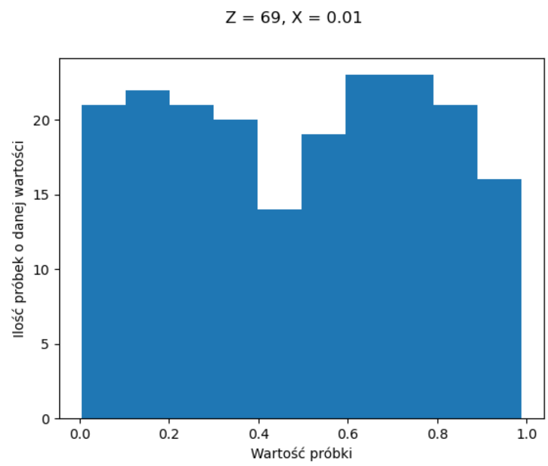
Tabela 1: Okres generatora w zależności od wartości Z

1.4 Okres generatora dla wybranych wartości Z

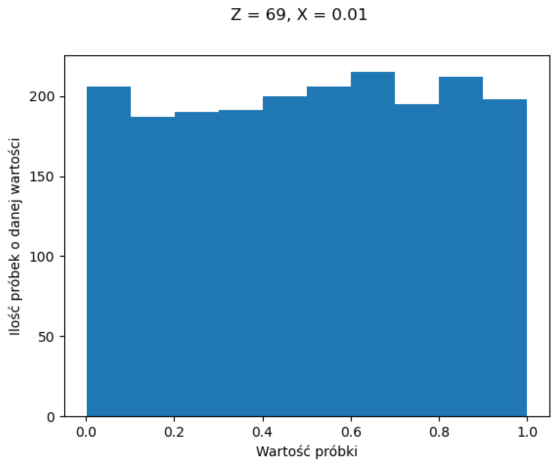
1.5 Podobieństwo histogramu ciągu wygenerowaych liczb, a gęstość rozkładu jednostajnego



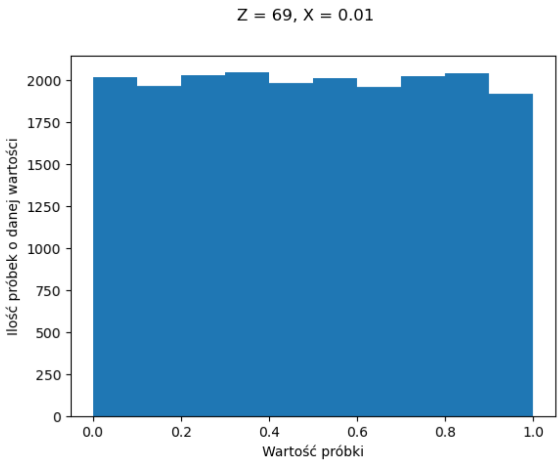
Rysunek 1: Ilość wygenerowanych próbek - 20



Rysunek 2: Ilość wygenerowanych próbek - 200



Rysunek 3: Ilość wygenerowanych próbek - 2000



Rysunek 4: Ilość wygenerowanych próbek - 20000

- Im więcej wygenerowaych próbek tym mocniej histogram upodabnia się do gęstości prawdopodobieństwa rozkładu jednostajnego.

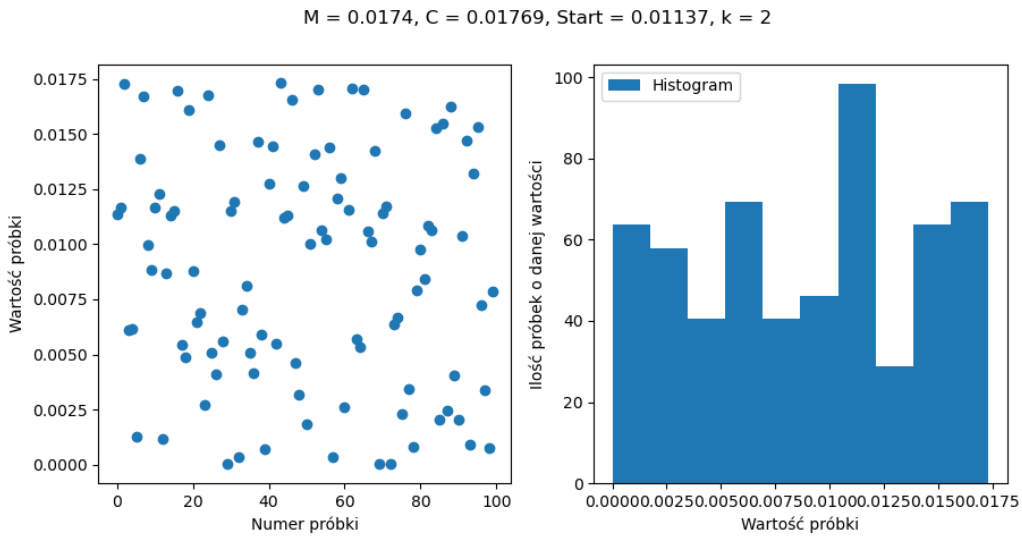
2 Generator dany równaniem

2.1 Opis

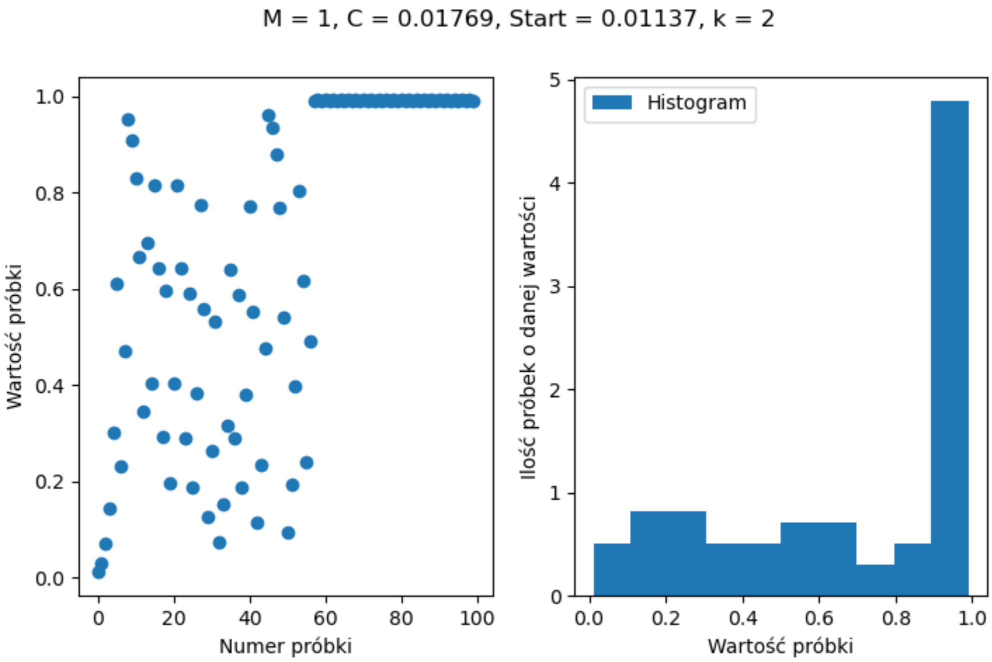
Zadanie polega na implementacji generatora liczb pseudolosowych oraz analizie wyników uzyskanych z jego udziałem. Generator oparty jest na równaniu:
$X_{n+1} = (a_0 X_n + a_1 X_{n-1} + \ldots + a_k X_{n-k} + C) \bmod m$

2.2 Zależność od współczynnika m

Współczynnik m odpowiada za zakres generowanych wartości.



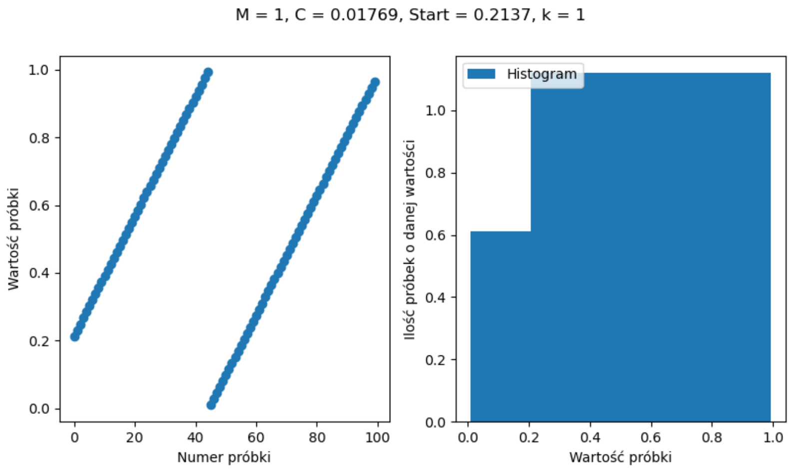
Rysunek 5: Zakres generowania liczb [0, 0.0174]



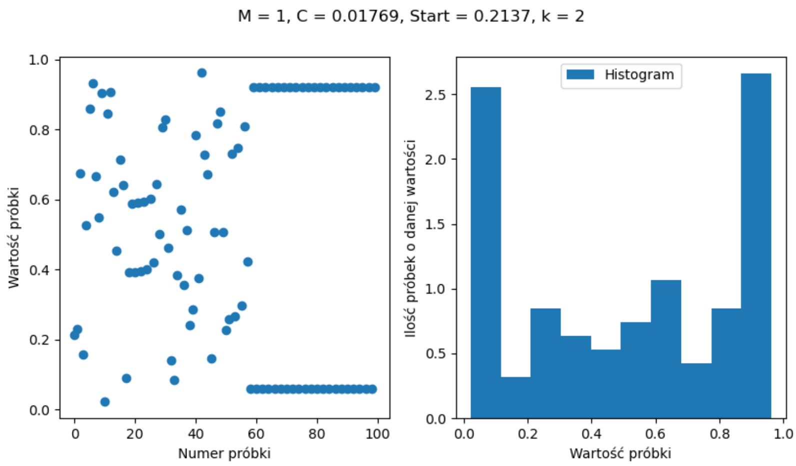
Rysunek 6: Zakres generowania liczb [0, 1]

2.3 Zależność od współczynnika k

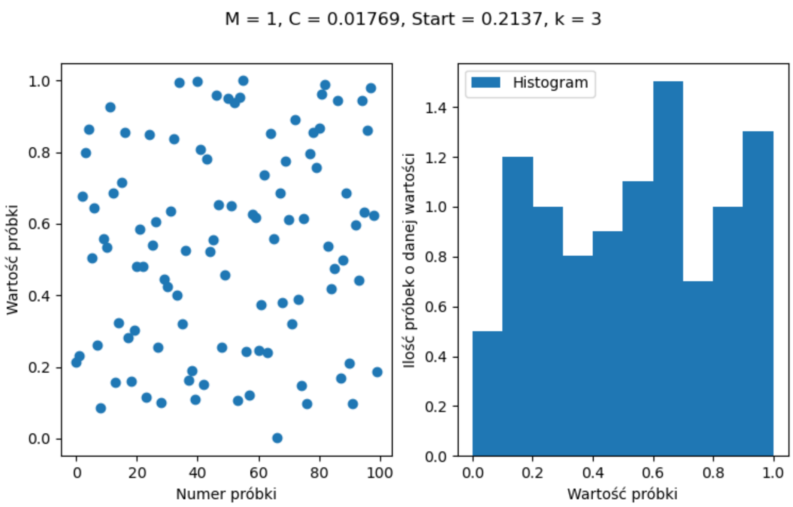
Współczynnik k odpowiada za ilość poprzednich próbek używanych podczas generowania nowych wartości.



Rysunek 7: Współczynnik k=1



Rysunek 8: Współczynnik k=2



Rysunek 9: Współczynnik k=3

- Współczynnik m odpowiada za zakres generowanych wartości. Wybór liczby całkowitej może spowodować zatrzymanie generatora.
- Ustawienie współczynnika k na wartość równą jeden uniemożliwia generowanie wartości losowych.
- Im większa wartość współczynnika k tym lepsze wyniki generatora.

3 Metoda odwracania dystrybuanty

3.1 Opis

Metoda odwracania dystrybuanty polega na odwróceniu funkcji dystrybuanty. Do uzyskania funkcji dystrybuanty posłużymy się całkowaniem funkcji opisującej gęstość prawdopodobieństwa analizowanego rozkładu.

3.2 Rozkład numer 1

Równanie funkcji rozkładu gęstość prawdopodobieństwa:

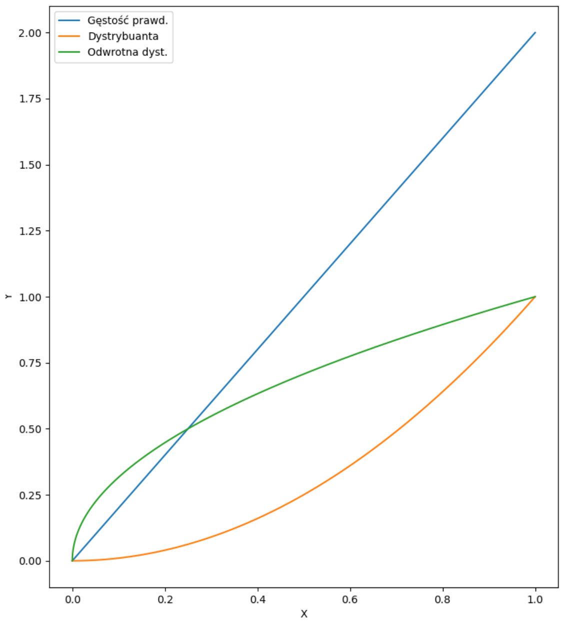
$$f(x) = \begin{cases} 2x \quad dla \quad x \in [0,1] \\ 0 \quad dla \quad x \in (-\infty, 0) \cup (1, \infty) \end{cases} \tag{1}$$

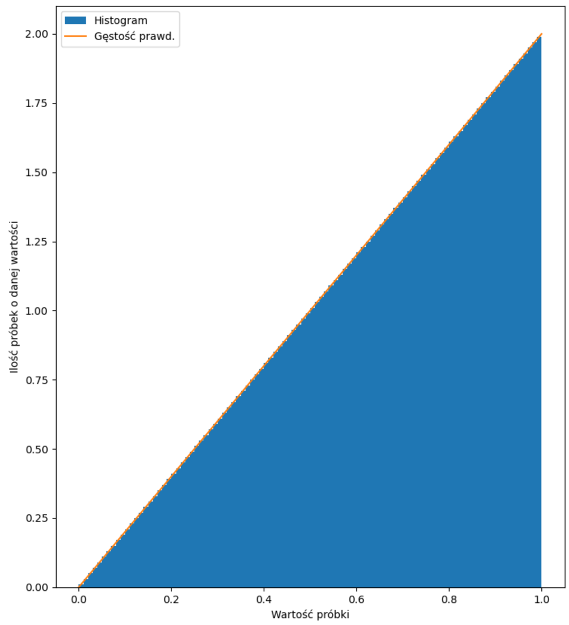
Dystrybuanta:

$$F(x) = \begin{cases} 0 \quad dla \quad x \in (-\infty, 0) \\ x^2 \quad dla \quad x \in [0,1] \\ 1 \quad dla \quad x \in (1, \infty) \end{cases} \tag{2}$$

Odwrotna dystrybuanta:

$$F^{-1}(y) = \sqrt{y}. \tag{3}$$





3.3 Rozkład numer 2

Równanie funkcji rozkładu gęstość prawdopodobieństwa:

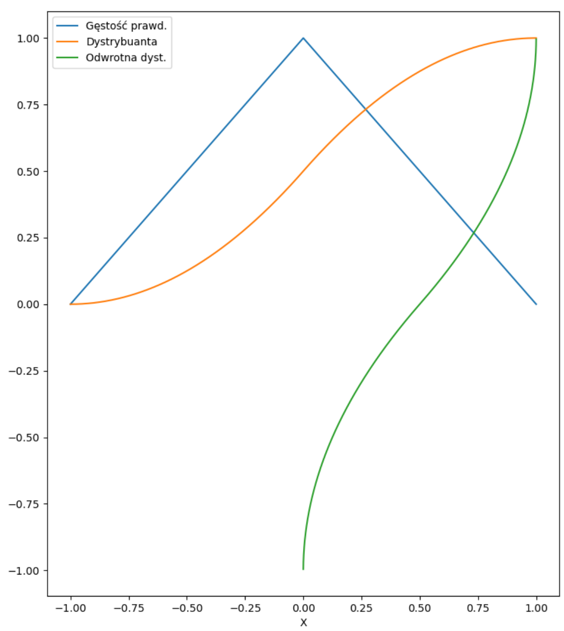
$$f(x) = \begin{cases} x+1 \quad dla \quad x \in (-1,0) \\ -x+1 \quad dla \quad x \in [0,1) \\ 0 \quad dla \quad x \notin (-1,1) \end{cases} \tag{4}$$

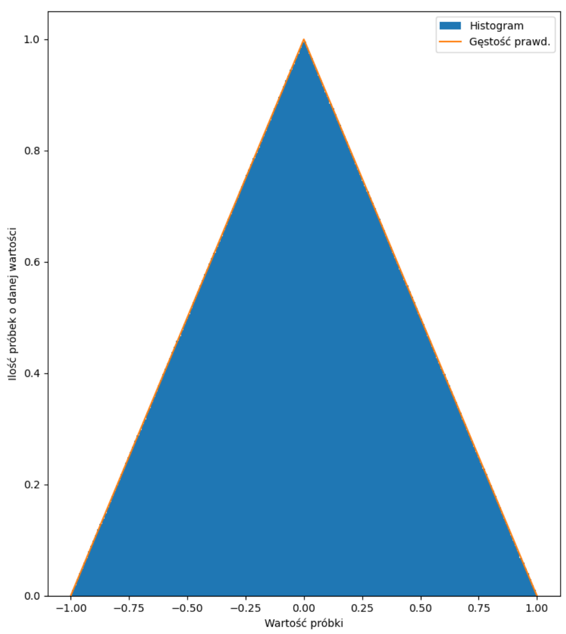
Dystrybuanta:

$$F(x) = \begin{cases} \frac{1}{2} - \frac{x^2}{2} + x \quad dla \quad x \in [0,1) \\ \frac{1}{2} + \frac{x^2}{2} + x \quad dla \quad x \in (-1,0) \\ 0 \quad dla \quad x \in (-\infty,-1] \\ 1 \quad dla \quad x \in [1,\infty) \end{cases} \tag{5}$$

Odwrotna dystrybuanta:

$$F^{-1}(y) = \begin{cases} \sqrt{2y} - 1 \quad dla \quad x \in [0,\frac{1}{2}] \\ 1 - \sqrt{2-2y} \quad dla \quad x \in (\frac{1}{2},1] \end{cases} \tag{6}$$





3.4 Rozkład wykładniczy

Równanie funkcji rozkładu gęstość prawdopodobieństwa:

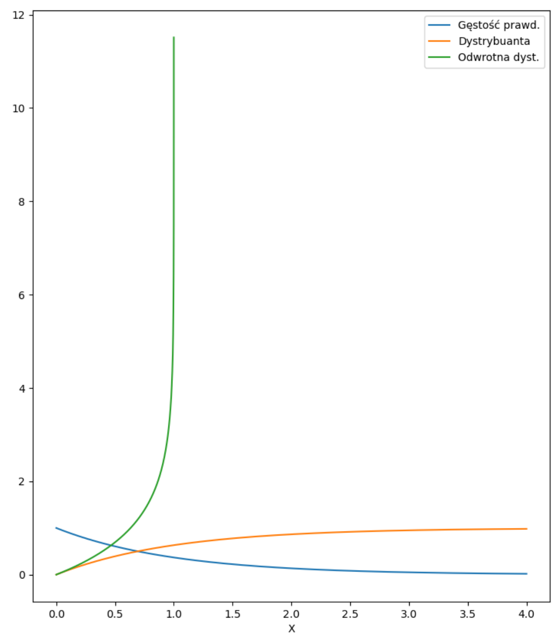
$$f(x) = e^{-x} \quad dla \quad x \in [0,\infty) \tag{7}$$

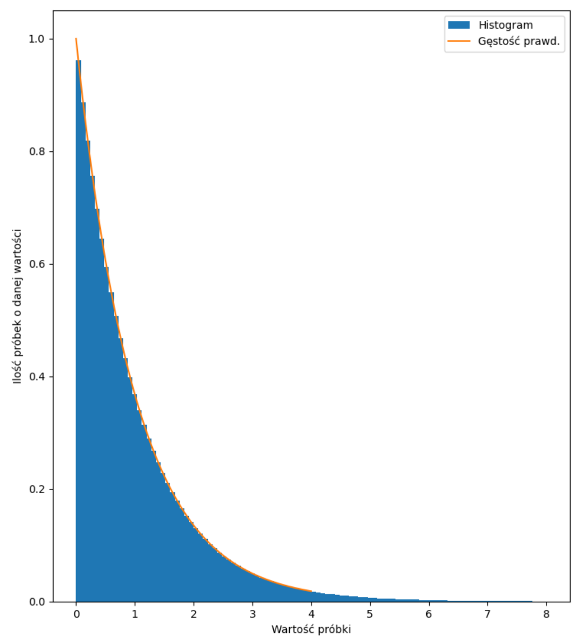
Dystrybuanta:

$$F(x) = 1 - e^{-x} \quad dla \quad x \in [0,\infty) \tag{8}$$

Odwrotna dystrybuanta:

$$F^{-1}(y) = -ln(1-y) \quad dla \quad x \in [0,\infty) \tag{9}$$





3.5 Rozkład Laplace'a

Równanie funkcji rozkładu gęstość prawdopodobieństwa:

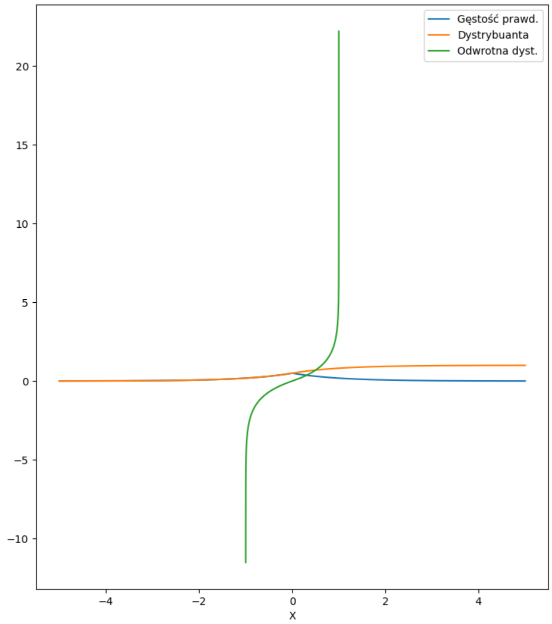
$$F(x) = \frac{1}{2}e^{-|x|} \tag{10}$$

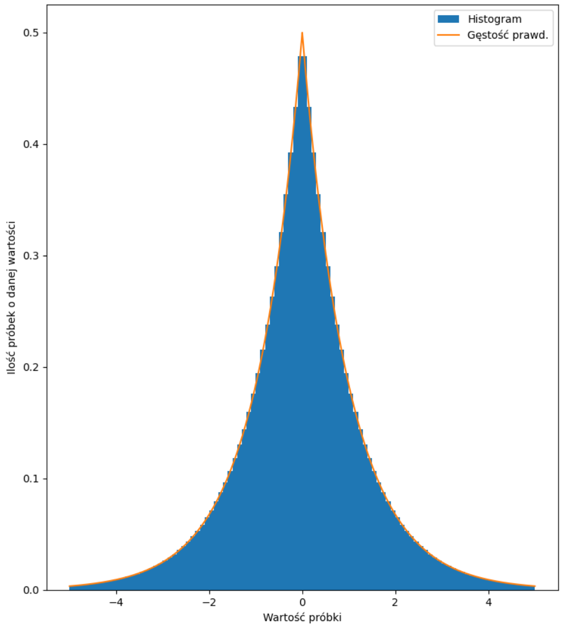
Dystrybuanta:

$$F^{-1}(y) = \begin{cases} \frac{1}{2} + \frac{1}{2}(1 - e^{-x}) & dla \quad x \in [0, \infty) \\ \frac{1}{2} - \frac{1}{2}(1 - e^{-x}) & dla \quad x \in (-\infty, 0) \end{cases} \tag{11}$$

Odwrotna dystrybuanta:

$$F^{-1}(y) = \begin{cases} -ln(1 - y) \quad dla \quad x \in [0, \infty) \\ ln(1 + y) \quad dla \quad x \in (-\infty, 0) \end{cases} \tag{12}$$





4 Metoda odrzucania

4.1 Rozkład numer 1

Równanie funkcji rozkładu gęstość prawdopodobieństwa:

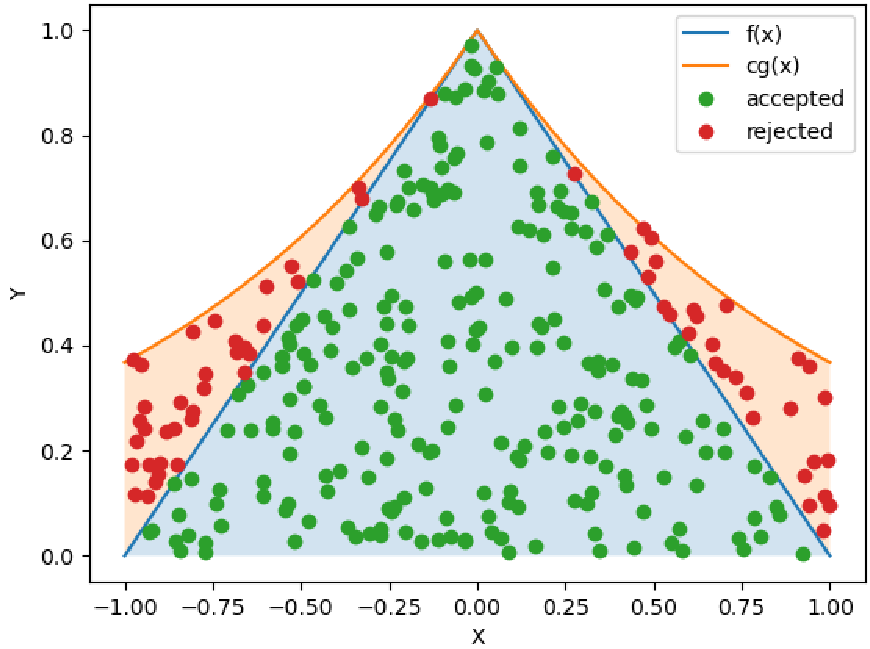
$$f(x) = \begin{cases} x+1 \quad dla \quad x \in (-1,0] \\ -x+1 \quad dla \quad x \in (0,1] \\ 0 \quad dla \quad x \in (-\infty,-1] \cup (1,\infty) \end{cases} \tag{13}$$

Przybliżenie funkcji f(x) z użyciem funkcji g(x):

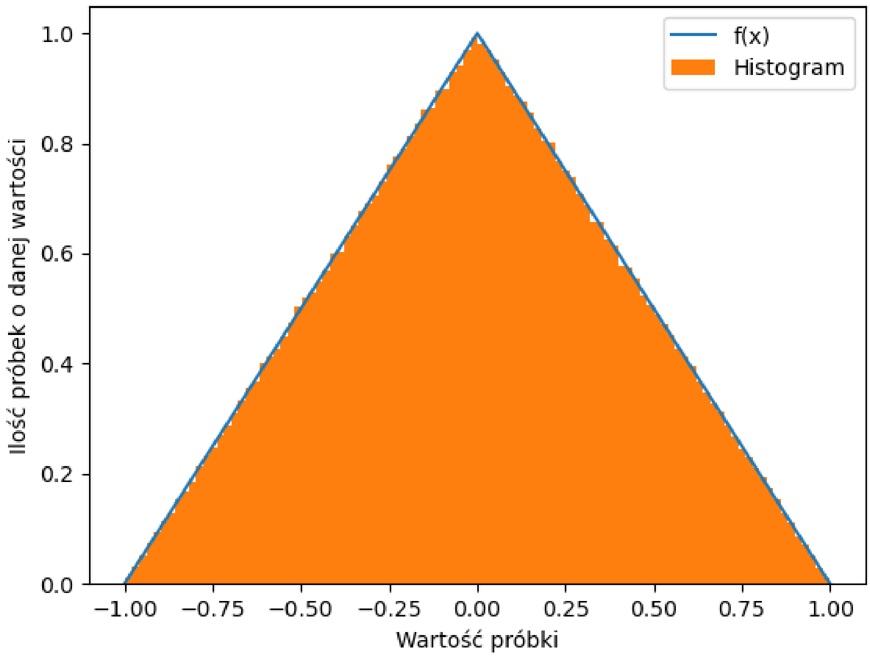
$$g(x) = \frac{1}{2}e^{-|x|} \tag{14}$$

Odwrotna dystrybuanta:

$$G^{-1}(y) = \quad dla \quad x \in [0,\infty) \begin{cases} ln(2x) \quad dla \quad x \in (-\infty, \frac{1}{2}] \\ -ln(2-2x) \quad dla \quad x \in (\frac{1}{2},\infty) \end{cases} \tag{15}$$



Rysunek 10: Generacja próbek rozkładu 1



Rysunek 11: Histogram rozkładu 1

4.2 Rozkład numer 2

Równanie funkcji rozkładu gęstość prawdopodobieństwa:

$$f(x) = \begin{cases} 50 \quad dla \quad x \in (0, \frac{1}{100}] \\ c \quad dla \quad x \in (\frac{1}{100}, 1] \end{cases} \tag{16}$$

Współczynnik c został wyznaczony w następujący sposób:

$$50 \cdot \frac{1}{100} + c \cdot 1 - \frac{1}{100} = 1 \tag{17}$$

$$c \cdot 1 - \frac{1}{100} = 1 - \frac{50}{100} \tag{18}$$

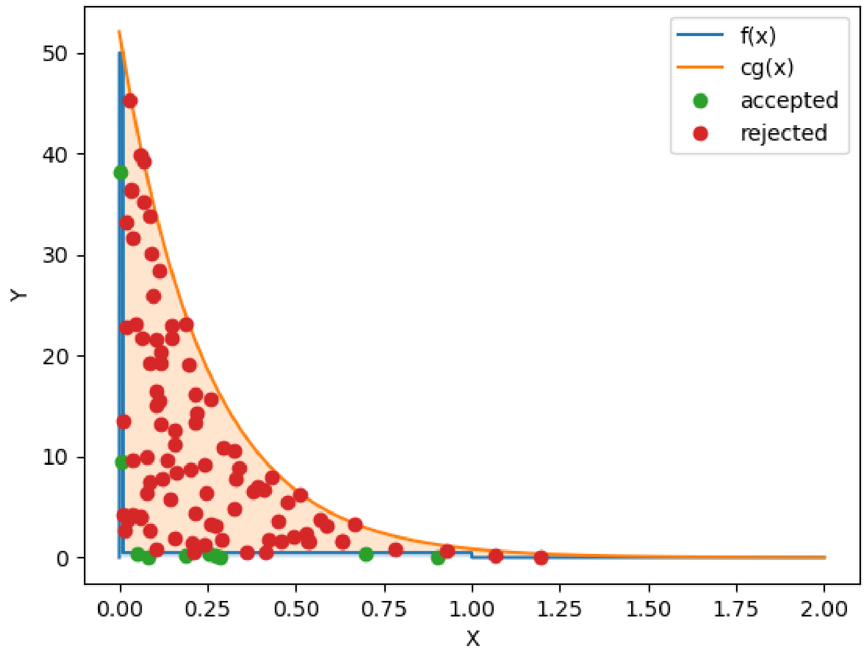
$$c = \frac{0.5}{\frac{99}{100}} \tag{19}$$

Przybliżenie funkcji f(x) z użyciem funkcji g(x):

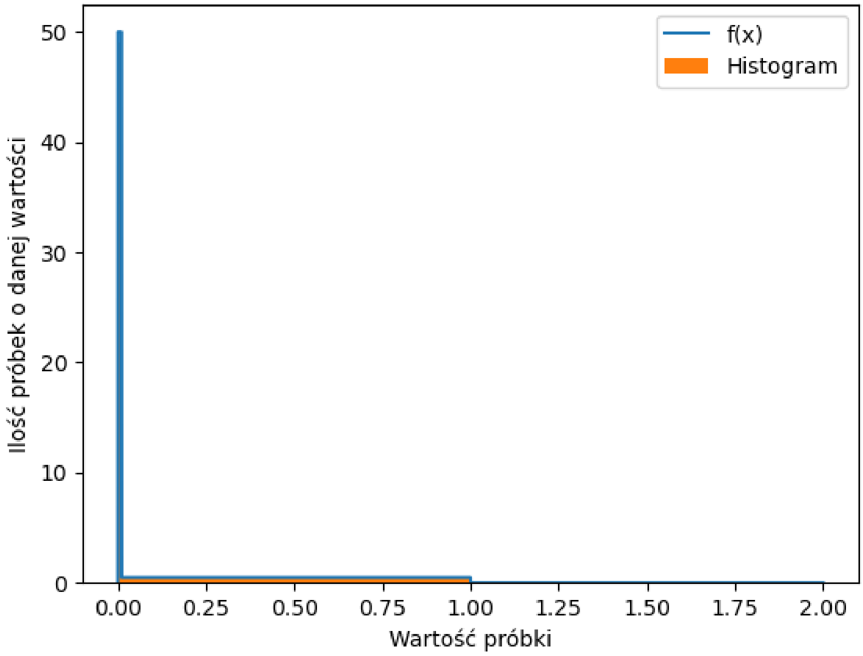
$$g(x) = 50 \cdot e^{-4.1 \cdot (x - 0.01)} \quad dla \quad x \in [0, \infty) \tag{20}$$

Odwrotna dystrybuanta:

$$g(x) = -0.244 \cdot ln(0.01919 \cdot x) \tag{21}$$



Rysunek 12: Generacja próbek rozkładu 2



Rysunek 13: Histogram rozkładu 2

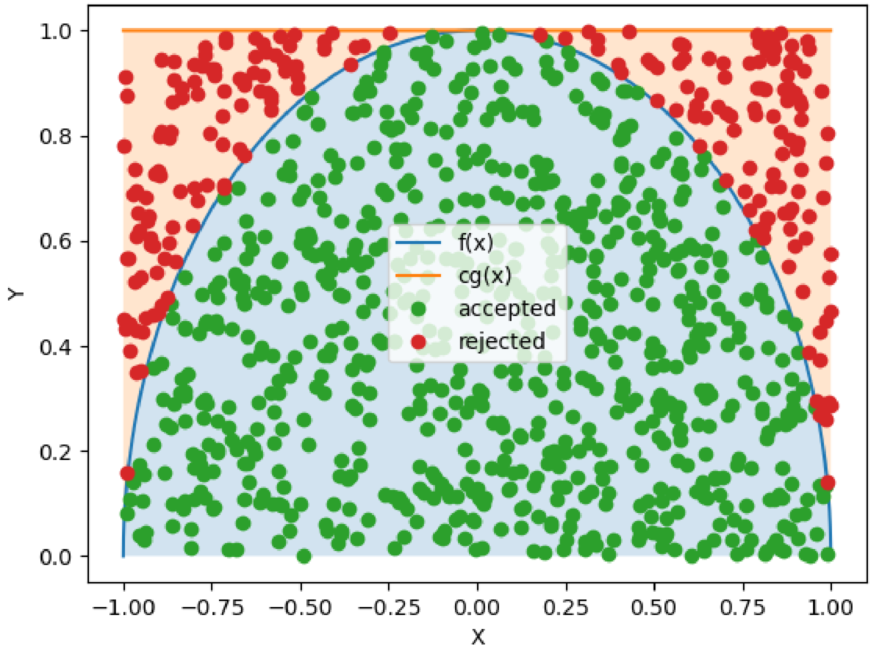
4.3 Półokrąg

Równanie funkcji rozkładu gęstość prawdopodobieństwa:

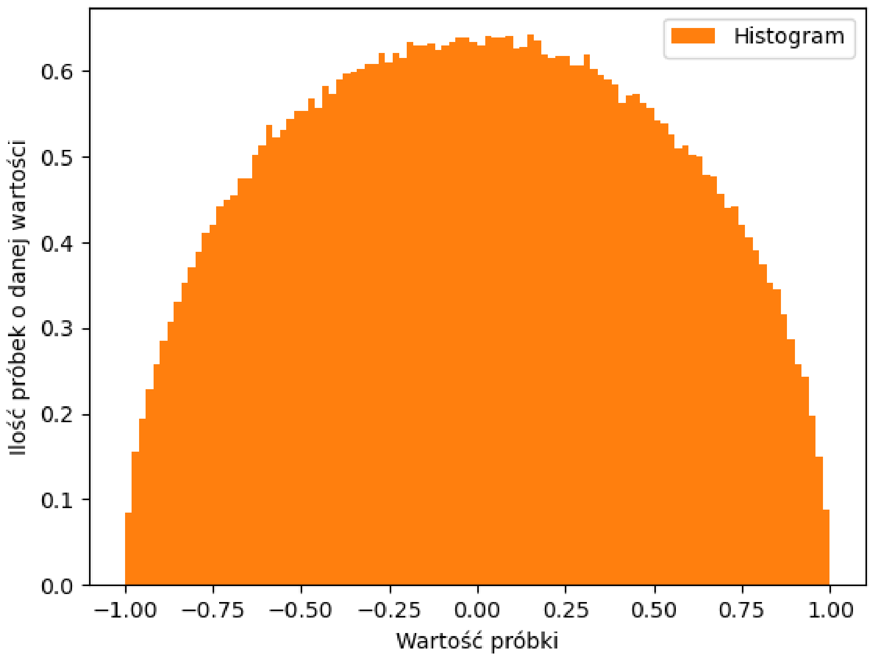
$$f(x) = \sqrt{r^2 - x^2} \tag{22}$$

Przybliżenie funkcji f(x) z użyciem funkcji g(x):

$$g(x) = 1 \quad dla \quad x \in [-1, 1] \tag{23}$$



Rysunek 14: Generacja próbek dla półokręgu



Rysunek 15: Histogram rozkładu dla półokręgu

4.4 Rozkład Normalny

Równanie funkcji rozkładu gęstość prawdopodobieństwa:

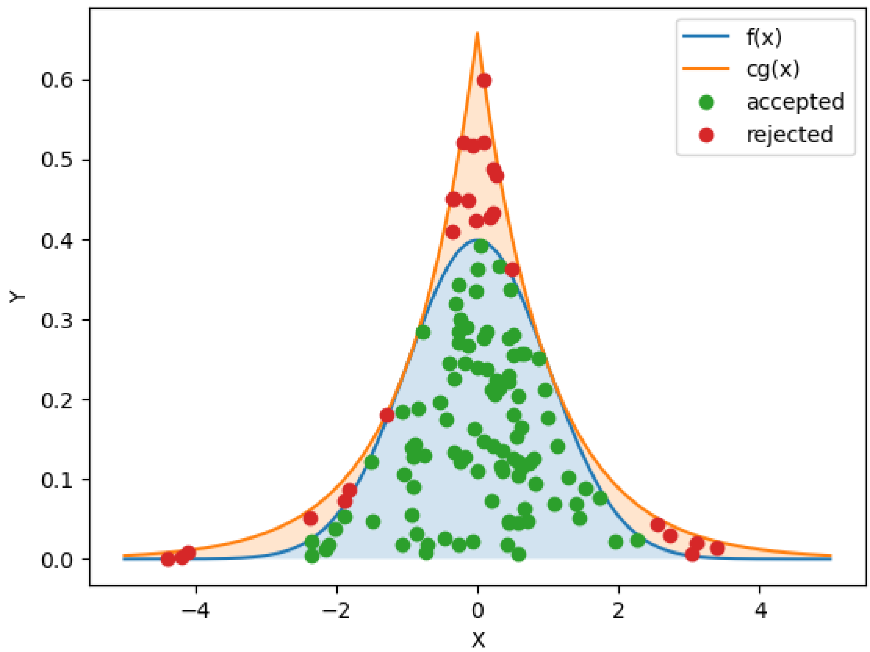
$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-x^2}{2}} \tag{24}$$

Przybliżenie funkcji f(x) z użyciem funkcji g(x):

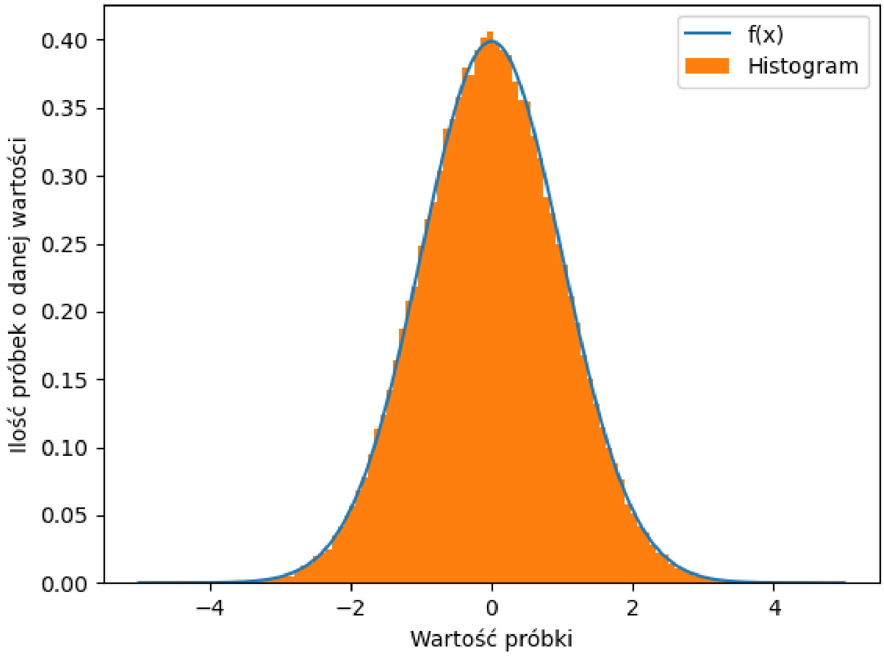
$$g(x) = \sqrt{\frac{2e}{\pi}} \cdot \frac{1}{2} \cdot e^{-|x|} \tag{25}$$

Odwrotna dystrybuanta:

$$g(x) = \begin{cases} ln(2x) \quad dla \quad x \in (-\infty, \frac{1}{2}] \\ -ln(2+2x) \quad dla \quad x \in (\frac{1}{2}, \infty) \end{cases} \tag{26}$$



Rysunek 16: Generacja próbek rozkładu normalnego



Rysunek 17: Histogram rozkładu normalnego

4.5 Wnioski

- Metoda eliminacji jest metodą stratna ponieważ w zależności od dobranej funkcji g(x), odrzucana jest pewna ilość losowań.
- Im lepiej dobierzemy funkcję pomocniczą tym większa efektywność generatora.
- Funkcja pomocnicza musi być odwrócona, więc należy dobierać funkcje możliwie trywialne.
- Stała C może być dobrana dowolnie duża, ale wyznaczenie jej analitycznie minimalizuje straty.

5 Estymacja

6 Podsumowanie