

AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY

WYDZIAŁ FIZYKI I INFORMATYKI STOSOWANEJ
KIERUNEK STUDIÓW: FIZYKA TECHNICZNA



METODY MONTE CARLO

Laboratorium 3

Generatory liczb pseudolosowych - rozkłady skorelowane w 2D

zrealizował
Przemysław Ryś

Kraków, 11 Marzec 2024

1 Opis zagadnienia

Wstęp

Rozkłady dwuwymiarowe są istotnym narzędziem w analizie danych oraz modelowaniu różnorodnych zjawisk z obszarów takich jak statystyka, matematyka finansowa czy przetwarzanie sygnałów. W niniejszym sprawozdaniu skupimy się na kilku kluczowych aspektach tych rozkładów, zwracając szczególną uwagę na rozkłady sferycznie konturowane, jednorodne w kole oraz transformacje afiniczne.

1.1 Rozkład sferycznie konturowany - normalny

Liczby pseudolosowe o rozkładzie normalnym generujemy metodą Boxa-Mullera. Algorytm dla rozkładu dwuwymiarowego można zapisać jako:

$$\begin{aligned}U_1 &\sim U(0, 1), \\U_2 &\sim U(0, 1), \\X &= \sqrt{-2 \ln(1 - U_1)} \cos(2\pi U_2), \\Y &= \sqrt{-2 \ln(1 - U_1)} \sin(2\pi U_2).\end{aligned}$$

Wektory (X, Y) mają rozkład sferycznie konturowany, ponieważ ich gęstość zależy tylko od odległości od środka rozkładu: $f(x, y) = f(x)f(y) = f(r)$, gdzie $r = \sqrt{x^2 + y^2}$.

1.2 Rozkład jednorodny w kole $K^2(0, 1)$

Dysponując rozkładem sferycznie konturowanym, możemy teraz umieścić wylosowane punkty na obwodzie okręgu o promieniu jednostkowym, normalizując zmienne:

$$\begin{aligned}X' &= \frac{X}{\sqrt{X^2 + Y^2}}, \\Y' &= \frac{Y}{\sqrt{X^2 + Y^2}}.\end{aligned}$$

Aby rozkład w kole był jednorodny, skalujemy zmienne zmienną losową z rozkładu o gęstości $h(r) = kr^{k-1}$, gdzie $k = 2$ to liczba wymiarów w naszym problemie.

1.3 Transformacja afiniczna: koło \rightarrow elipsa

Rozkład dwuwymiarowy możemy podać transformacji liniowej (afinicznej), która przekształci koło w elipsę. Docelowy kształt elipsy definiujemy podając wektory określające półosie główne:

$$\vec{r}_1 = \begin{pmatrix} r_{11} \\ r_{21} \end{pmatrix}, \quad \vec{r}_2 = \begin{pmatrix} r_{12} \\ r_{22} \end{pmatrix}.$$

Wektory te stanowią kolumny macierzy transformacji $A = [\vec{r}_1 | \vec{r}_2]$. Macierz A określa obrót oraz skalowanie wzdłuż pól głównych, a rozkład możemy przesunąć o wektor $\vec{c}^T = [c_1, c_2]$. Transformację możemy zapisać w postaci macierzowej:

$$\vec{r}' = A\vec{r} + \vec{c}.$$

1.3.1 Wybór osi i skalowanie

Ponieważ półosie główne elipsy muszą być ortogonalne, wystarczy obrócić je o zadany kąt α przy użyciu macierzy obrotu R_α i przeskalować ich długości:

$$\vec{r}_1 = b_1 R_\alpha \vec{b}_x, \quad \vec{r}_2 = b_2 R_\alpha \vec{b}_y,$$

gdzie $\vec{b}_x = [1, 0]^T$, $\vec{b}_y = [0, 1]^T$, a macierz obrotu (przeciwnie do ruchu wskazówek zegara) R_α ma postać:

$$R_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Współczynniki skalujące b_1, b_2 są większe od zera.

1.4 Wyznaczanie macierzy kowariancji

Macierz kowariancji Σ dla rozkładu dwuwymiarowego ma postać:

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{pmatrix}.$$

Jej elementy liczymy następująco:

$$\begin{aligned} \sigma_x^2 &= \int \int (x - \mu_x)^2 f(x, y) dx dy \\ &= \int \int x^2 f(x, y) dx dy - \mu_x^2 = \langle x^2 \rangle - \langle x \rangle^2, \\ \sigma_{xy} &= \sigma_{yx} = \int \int (x - \mu_x)(y - \mu_y) f(x, y) dx dy \\ &= \int \int xy f(x, y) dx dy - \mu_x \mu_y = \langle xy \rangle - \langle x \rangle \langle y \rangle. \end{aligned}$$

Elementy macierzy kowariancji możemy wykorzystać do wyznaczenia współczynnika korelacji zmiennych:

$$r_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2 \sigma_y^2}}, \quad r \in [-1, 1].$$

1.5 Transformacja afiniczna a macierz kowariancji dla rozkładu gaussowskiego

Ogólna postać funkcji gęstości prawdopodobieństwa dla k-wymiarowego rozkładu normalnego $N_k(0, 1)$:

$$f(\vec{r}) = \frac{1}{\sqrt{(2\pi)^k |D|}} \exp\left(-\frac{1}{2} \vec{r}^T D^{-1} \vec{r}\right),$$

gdzie D to macierz kowariancji. Transformacja afiniczna zmiennych ma postać:

$$\vec{r}' = A\vec{r}, \quad \vec{r} = A^{-1}\vec{r}',$$

co po podstawieniu do funkcji gęstości prawdopodobieństwa daje zmianę w wykładniku:

$$\vec{r}^T D^{-1} \vec{r} = \vec{r}'^T (A^{-1})^T D^{-1} A^{-1} \vec{r}' = \vec{r}'^T \Sigma^{-1} \vec{r}'.$$

Nowa macierz kowariancji dla pierwotnego rozkładu $N_2(0, 1)$:

$$\begin{aligned} \Sigma &= A D A^T, \\ D &= \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} = I, \\ \Rightarrow \Sigma &= A A^T. \end{aligned}$$

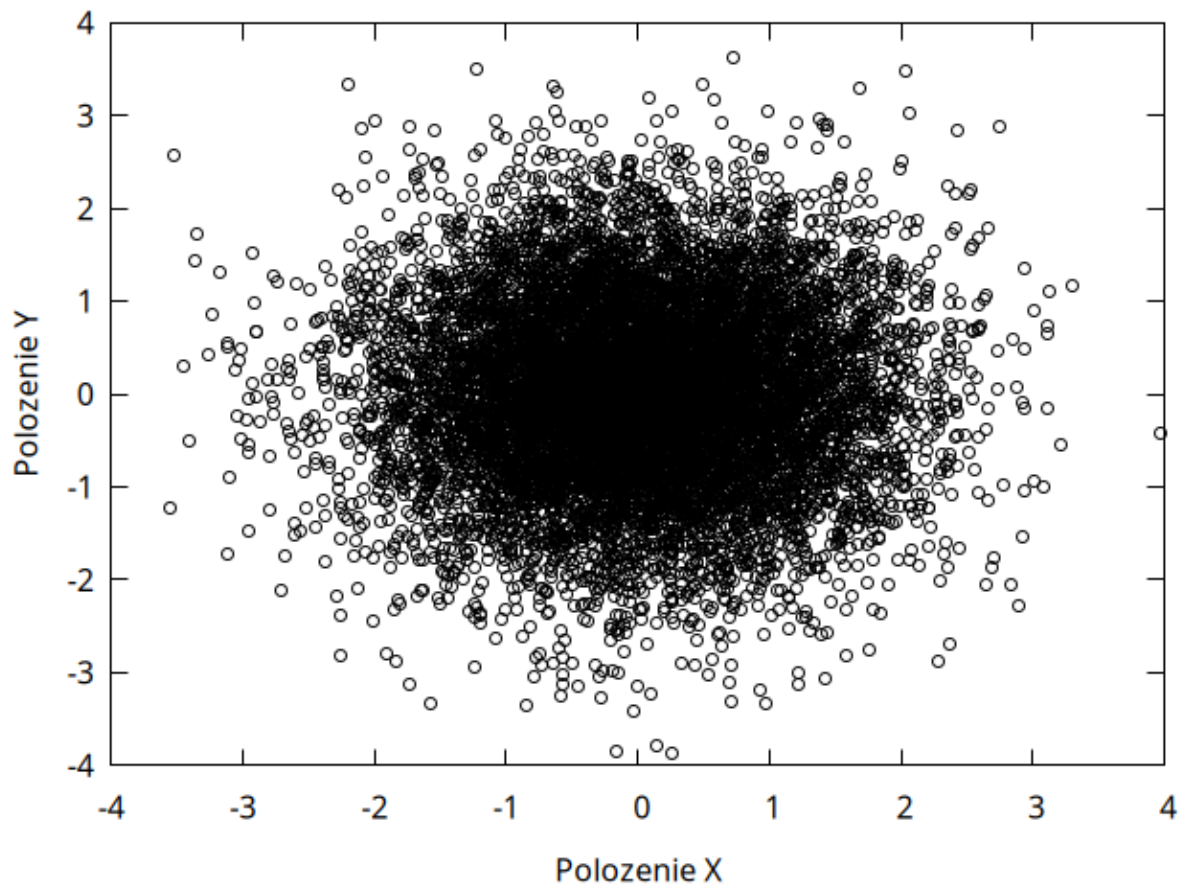
Jeśli oznaczmy $A^{-1}\vec{r} = \vec{z}$, to losując wektory \vec{z} z rozkładu $N_2(0, 1)$ dostaniemy rozkład skorelowany:

$$\vec{r}' = A\vec{z},$$

określony macierzą kowariancji $\Sigma = A A^T$.

2 Wyniki

Zadanie 1



Rys. 1: Wykres położenia punktów wylosowanych z dwuwymiarowego rozkładu normalnego $N^2(0, 1)$ metodą Boxa-Mullera

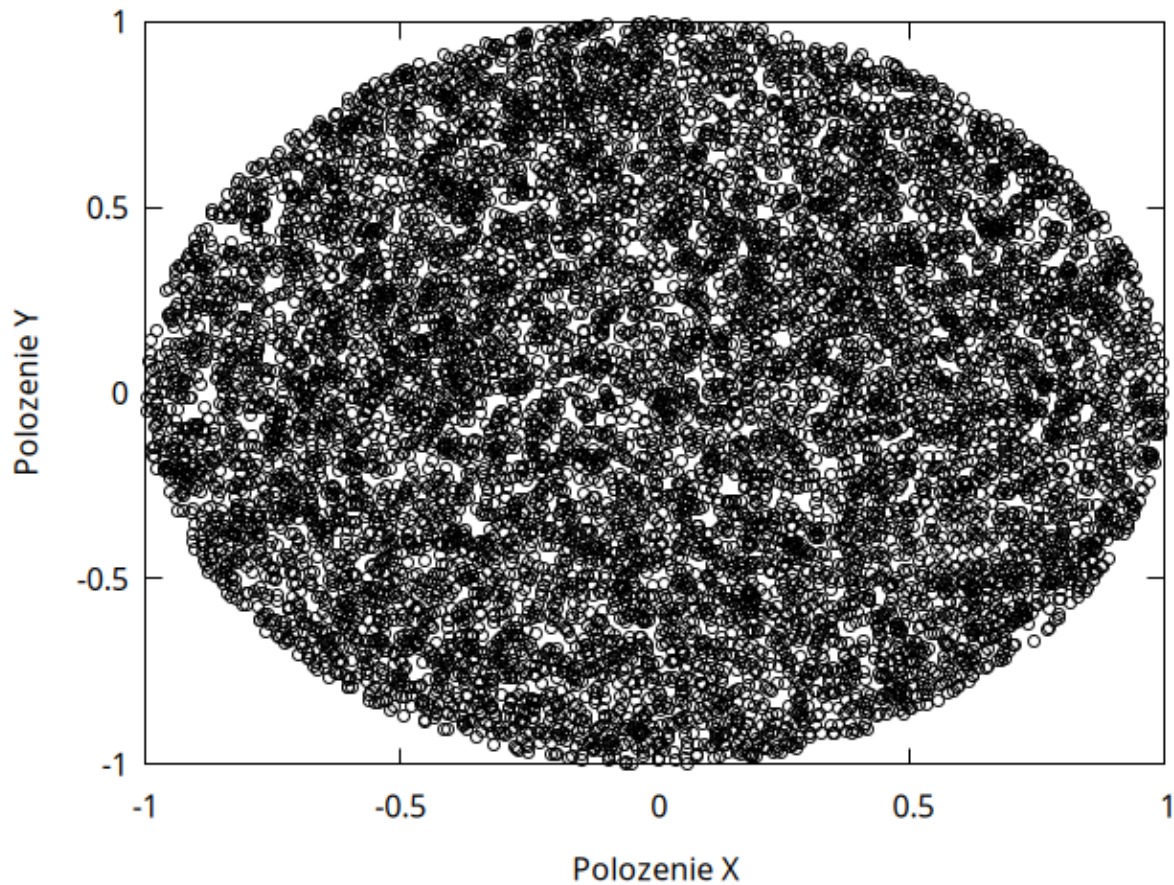
Wylosowano $n = 10^4$ punktów z dwuwymiarowego rozkładu normalnego $N^2(0, 1)$ przy użyciu metody Boxa-Mullera. Rysunek 1 przedstawia wylosowane punkty tym rozkładem, można na nim zauważyć szczególną koncentrację punktów w pobliżu centrum wykresu, co sugeruje że dwuwymiarowy rozkład normalny został zrealizowany poprawnie.

Zadanie 2



Rys. 2: Wykres położenia punktów wylosowanych wewnątrz koła o promieniu jednostkowym przy wykorzystaniu rozkładu sferycznie konturowanego

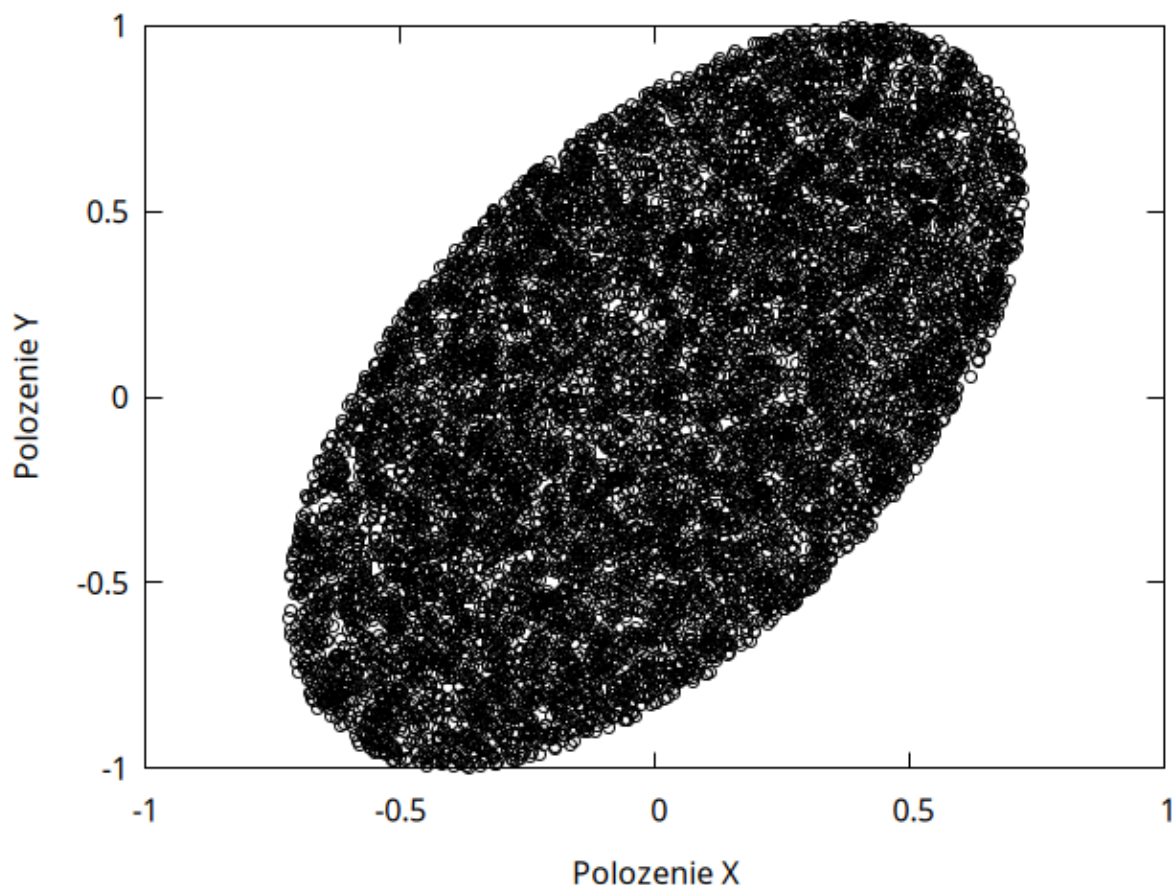
Wygenerowane z poprzedniego zadania punkty poddano transformacji opisanej w sekcji 1 w wyniku czego punkty rozmieszczone zostały na obwodzie okręgu o promieniu jednostkowym. Jak widzimy na rysunku 2 rzeczywiście wszystkie wylosowane punkty przemieściły się na wspomniany obwód. Jest to użyteczne gdy w zagadnieniu dla którego losujemy dane mamy narzucony warunek, by losować dane z odpowiedniej powierzchni, co jest łatwo rozszerzalne dla większych wymiarów.



Rys. 3: Wykres położeń punktów wylosowanych wewnątrz koła o promieniu jednostkowym przy wykorzystaniu rozkładu sferycznie konturowanego

W kolejnym kroku punkty obwodu zostały przekształcone poprzez skalowanie ich zmienną losową postaci $h(r) = 2r = 2\sqrt{x^2 + y^2}$ w celu jednorodnego rozmieszczenia danych wewnątrz zadanego koła jednostkowego. Realizację tego zadania zobrazowano na rysunku 3. Widzimy że koncentracja bliżej centrum nie tyle się zmniejszyła, co jest w przybliżeniu taka sama w każdym punkcie koła.

Zadanie 3



Rys. 4: Wykres położenia punktów przetransformowanych macierzą obrotu A

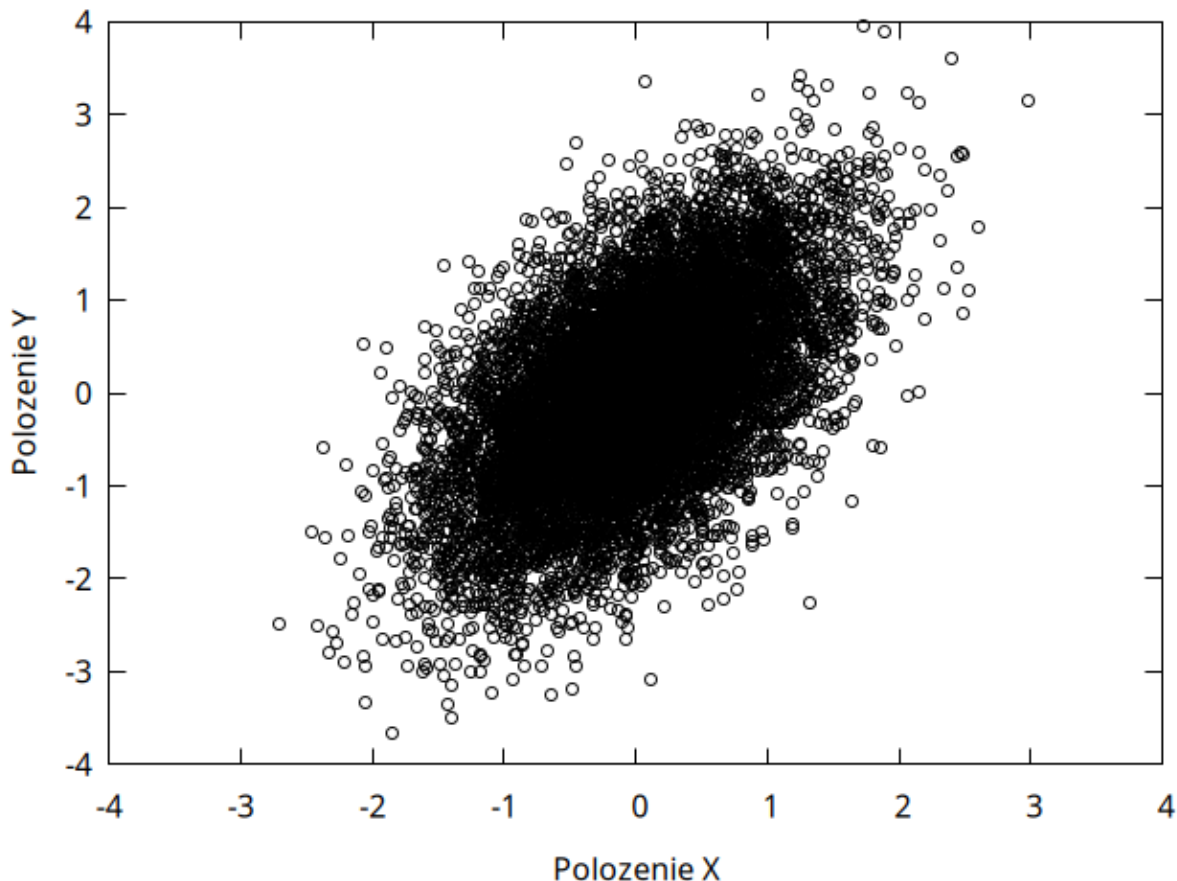
Kolejnym procesem jakiemu poddano dane była ich transformacja macierzą, której kolumny stanowiły wektory poddane transformacji afinicznej. odpowiednie wersory osi wyjściowego układu kartezjańskiego zostały obrócone o kąt $\frac{\pi}{4}$ przeciwnie do ruchu wskazówek zegara, oraz przeskalowane odpowiednio wersor osi x przez wartość 1, natomiast y przez wartość 0.2. Efektywnie transformacja taką macierzą działająca na punkty układające się we wzór koła przetransformowywane są na odpowiadające punkty elipsy. Zostało to zestawione na obrazie 4. ich rozkład we wnętrzu jest w dalszym ciągu jednorodny.

W zrealizowanym programie obliczono również macierz kowariancji efektywnie uzyskanych punktów, które uległy transformacji do wnętrza sfery, ponieważ są określone rozkładem jednorodnym. Macierz ta wynosi:

$$\begin{bmatrix} 0.130132 & 0.100454 \\ 0.100454 & 0.25255 \end{bmatrix}$$

Na jej podstawie obliczono wartość współczynnika korelacji, który wyniósł: $r_{xy} = 0.554115$. Wartość ta leży mniej więcej po środku pomiędzy 0 oznaczającą brak korelacji, a wartością 1(-1) oznaczającą pełną korelację dodatnią(ujemną) w związku z czym nie można jednostecznie określić zależności pomiędzy położeniem x oraz y przekształconych do elipsy punktów rozkładu jednorodnego.

Zadanie 4



Rys. 5: Wykres położenia punktów przetransformowanych macierzą obrotu A

Analogicznie jak zostało to zrealizowane dla punktów rozkładu jednorodnego w kole jednostkowym. wykonano transformacje danych na nowo wylosowanych z dwuwymiarowego rozkładu normalnego metodą Boxa-Mullera, tak jak zostało to zrealizowane w zadaniu 1. Jak możemy zauważyć z rysunku 5 rozkład losowanych punktów również się nie zmienia i w dalszym ciągu jest normalny niemniej jest on realizowany w przekształconej afinicznie przestrzeni tj. na elipsie.

Tak jak poprzednio, zrealizowano również obliczanie macierzy kowariancji i dla przypadku punktów rozkładu normalnego ułożonych na uzyskanej po transformacji elipsie. Obliczona macierz kowariancji wynosi:

$$\begin{bmatrix} 0.519366 & 0.402261 \\ 0.402261 & 1.00711 \end{bmatrix}$$

Na jej podstawie obliczono wartość współczynnika korelacji, który wyniósł:
 $r_{xy} = 0.556201$. Wartość ta również jest po środku wartości 0 a 1, w związku z czym nie można jednostecznie określić zależności pomiędzy położeniem x oraz y przekształconych punktów rozkładu normalnego.

3 Wnioski

Podsumowując zagadnienie polegające na generowaniu liczb z rozkładów można zrealizować na różne sposoby. Przyjęcie jednej konkretnej metody zależy od wykonywanego zadania oraz charakteru samych zmiennych. Korzystając z narzędzi algebry liniowej, można to zrobić bardzo prosto, implementując odpowiednie przekształcenia można wygenerować dla własnych potrzeb rozkłady nie tylko o zadanym charakterze względem pozostałych danych, ale również zrealizować je na dowolnej powierzchni, czy też ich wnętrzu, w dodatku wychodząc od podstawowego generatora jednorodnego liczb z zakresu $(0,1)$.