Metody Monte Carlo - laboratorium 3

Temat: Generowanie wybranych rozkładów gęstości prawdopodobieństwa

Zadanie 1

Korzystając z RNG umieszczonych w bibliotece *gsl* (np. *Ranlux* lub *Mersennse Twister*) napisać program generujący liczby losowe zgodnie z rozkładem normalnym (Gaussa) $f(x)_{\mu,\sigma} = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$. W tym celu wykorzystać algorytm oparty o **Centralne Twierdzenie Graniczne** (poniżej w ramce):

> Niech X_i mają rozkład $U(0, 1)$, i=1, 2, ..., n oraz $R_n = \sum_{i=1}^{n} X_i$. Wówczas rozkład zmiennej $Y = \frac{R_n - n/2}{\sqrt{n/12}}$ dąży do standardowego rozkładu normalnego $N(0, 1)$. Praktyczny wybór $n = 12$
>
> prowadzi do $Y = R_{12} - 6$.

Następnie wyznaczyć histogram dla przedziału zmienności $x \in [-6, 6]$, wyświetlić go przy użyciu GNUPLOT i porównać z kształtem rzeczywistej funkcji gęstości prawdopodobieństwa $N(0, 1)$. Można ją wyświetnić wydając polecenia w środowisku GNUPLOT:

```
normal(x,mi,sigma)=1/(sigma*sqrt(2*pi))*exp(-(x-mi)**2/(2*sigma**2))
plot [-6:6] normal(x,0,1)
```

Zadanie 2

Korzystając z RNG umieszczonych w bibliotece *gsl* (np.: *Ranlux* lub *Mersennse Twister*) napisać program generujący liczby losowe zgodnie z rozkładem normalnym dwuwymiarowym $f(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$ w oparciu o metodę **Boxa-Mullera** (poniżej w ramce):

> Niech X_1 *i* X_2 mają rozkład $U(0, 1)$. Wtedy:
>
> $$x = \sqrt{-2\ln X_1}\cos(2\pi X_2)$$
> $$y = \sqrt{-2\ln X_1}\sin(2\pi X_2)$$
>
> są niezależnymi zmiennymi o rozkładach $N(0, 1)$.

Następnie wyznaczyć histogram, wyświetlić go przy użyciu GNUPLOT i porównać z kształtem rzeczywistej dwuwymiarowej funkcji gęstości prawdopodobieństwa *N*(0, 1). Można ją wyświetlić wydając polecenia w środowisku GNUPLOT:

```
normal_BM(x,y)=1/(2*pi)*exp(-(x**2+y**2)/2)
splot [-6:6][-6:6] [0:0.2] normal_BM(x,y) with fs
```

To samo wykonać w oparciu o algorytm **Marsaglii i Braya**:

Algorytm:

1) wygenerować A_1 *i* A_2 z rozkładu $U(0, 1)$ i obliczyć $X_1 = 2A_1 - 1$ i $X_2 = 2A_2 - 1$
2) obliczyć : $B = X_1^2 + X_2^2$
3) jeżeli $B > 1$, to powrót do punktu 1
4) wyznaczyć: $x = X_1 Z \;\; i \;\; y = X_2 Z \;\; gdzie \;\; Z = \sqrt{\frac{-2\ln B}{B}}$

Sprawdzić który z nich jest bardziej efektywny (np. mierząc czas generacji 10^7 par liczb)

Zadanie 3

Korzystając z **metody eliminacji** (przypomnienie w ramce) wygenerować liczby losowe o rozkładzie zdefiniowanym gęstością prawdopodobieństwa $f(x) = \frac{5}{12}[1+(x-1)^4], \;\; 0 \le x \le 2$, a następnie wyznaczyć histogram i porównać go z kształtem rzeczywistej funkcji gęstości prawdopodobieństwa.

> Metoda eliminacji
>
> 1) Generujemy dwie niezależne zmienne losowe U_1 i U_2 o rozkładach równomiernych odpowiednio $U(a, b)$ i $U(0, d)$, gdzie a–b to zmienność x, zaś 0–d to zmienność $f(x)$.
>
> 2) Jeżeli $U_2 < f(U_1)$, to przyjmujemy $X = U_1$. W przeciwnym przypadku parę liczb U_1, U_2 pomijamy i wracamy do punktu 1.

To samo wykonać wykorzystując **metodę superpozycji** w postaci: $f(x) = p_1 \cdot g_1(x) + p_2 \cdot g_2(x)$

gdzie $g_1(x) = \frac{1}{2}, \;\; g_2(x) = \frac{5}{2}(x-1)^4, \;\; p_1 = \frac{5}{6}, \;\; p_2 = \frac{1}{6}$ i ostatecznie:

<u>Metody Monte Carlo - laboratorium 3</u>

$$X=\begin{cases} 2U_2, & U_1 < 5/6 \\ 1+(2U_2-1)^{1/5}, & U_1 \geq 5/6 \end{cases} \quad U_1, U_2 \text{ mają rozkład } U(0,1)$$

Porównać efektywność obu algorytmów.

Uwaga! Należy "obejść" ograniczenie funkcji `pow(double, double)` z biblioteki matematycznej, która zwraca zero, gdy podstawa potęgi jest ujemna.

Zadanie 4

Korzystając z funkcji bibliotecznych *gsl* wygenerować liczby losowe o rozkładach gęstości prawdopodobieństwa: Laplace'a, Pareto, Gamma i Dirichleta.
W tym celu należy zapoznać się z dokumentacją biblioteki *gsl* znajdującą się w katalogu /**gsl** (plik *index.html*) w podrozdziale *Random Number Distributions* (szczególnie ciekawy: *Random Number Distribution Examples*).

Zadanie 5

Korzystając z podanych poniżej związków wygenerować rozkłady: Cauchy'ego i trójkątny.

1. Jeżeli zmienne X i Y są niezależnymi zmiennymi losowymi o rozkładach $N(0, 1)$, to zmienna $Z = X/Y$ ma rozkład Cauchy'ego $C(0, 1)$.
2. Jeśli X i Y są niezależnymi zmiennymi losowymi o rozkładzie równomiernym na przedziale $(a/2, b/2)$, to zmienna losowa $Z = X + Y$ ma rozkład trójkątny na przedziale (a, b).

Uzyskane histogramy wykreślić na tle funkcji analitycznych.