Generowanie ciągu liczb pseudolosowych o rozkładzie jednorodnym i trójkątnym.

Tomasz Chwiej

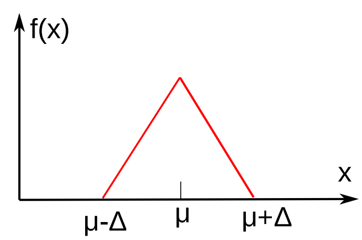
11 czerwca 2018

1 Wstęp

Funkcję gęstości prawdopodobieństwa dla rozkładu trójkątnego $T(\mu, \Delta)$ (rys.1) definiujemy następująco

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

gdzie: μ to środek rozkładu, a Δ to jego szerokość.



Rysunek 1: Funkcja gęstości prawdopodobieństwa rozkładu trójkątnego.

Dystrybuanta tego rozkładu jest następująca

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

Jeśli $\xi_1 \in U(0,1)$ i $\xi_2 \in U(0,1)$ to zmienną o rozkładzie trójkątnym oraz parametrach μ i Δ generujemy stosując formułę

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

2 Zadania do wykonania

2.1 Rozkład jednorodny

Startując od $x_0 = 10$ należy wygenerować $n = 10^4$ liczb pseudolosowych przy użyciu generatora mieszanego

$$x_{n+1} = (ax_n + \ c)\ mod\ m \tag{4}$$

o parametrach (**typu long int**):

a) $a = 123$, $c = 1$, $m = 2^{15}$

b) $a = 69069$, $c = 1$, $m = 2^{32}$

Proszę w obu przypadkach sporządzić rysunek $X_{i+1} = f(X_i)$ ($X_i = x_i/(m+1.0)$ **z warunku normalizacji do rozkładu U(0,1)**). Czy porównując oba rysunki można stwierdzić, który generator ma lepsze własności statystyczne? W sprawozdaniu proszę uzasadanić odpowiedź. W sprawozdaniu proszę także zamieścić histogram (dla $k = 12$ podprzedziałów) rozkładu gętości prawdopodobieństwa dla $n = 10^4$ liczb pseudolosowych o rozkładzie równomiernym (oba przypadki). Proszę także podać obliczone wartości μ i σ i porównać je z wartościami teoretycznymi. Uwaga: Dla generatorów proszę napisać funkcje w których zmienna x będzie typu **static long long int x=10;** tj. będzie ona inicjalizowana tylko podczas pierwszego wywołania a jej aktualna wartość będzie zachowywana w pamięci po zakończeniu działania funkcji.

```
double gen_1(){
   static long int x=10;
   int a=...;
   int c=...;
   long int m=...;
   x=(a*x+c) % m;
   return x/(m+1.0);
}
```

2.2 Rozkład trójkątny

1. Wygenerować $n = 10^3$ liczb o rozkładzie trójkatnym (wzór 3) o parametrach $\mu = 4$ i $\Delta = 3$.

2. Podzielić przedział $[\mu - \Delta, \mu + \Delta]$ na $K = 10$ podprzedziałów i zliczyć ile liczb wpada do każdego z nich.

3. Dla rozkładu trójkatnego przeprowadzić test χ^2 tj. określić wartość statystyki testowej

$$\chi^2 = \sum_{i=1}^{K} \frac{(n_i - n \cdot p_i)^2}{n \cdot p_i} \tag{5}$$

   gdzie: n_i to ilość liczb znajdujących się w podprzedziale o indeksie i, p_i to prawdopodobieństwo teoretyczne że zmienna losowa X znajdzie się w $i - tym$ przedziale

$$p_i = F(x_{i,max}) - F(x_{i,min}) \tag{6}$$

   We wzorze (6) $F(x)$ jest wartością dystrybuanty liczonej zgodnie z wzorem (2). Wartości: p_i oraz $n \cdot p_i$ dla każdego z podprzedziałów zapisać do pliku. W sprawozdaniu proszę zamieścić histogram pokazujący wartości n_i/n oraz p_i dla każdego z podprzedziałów.

4. Testujemy hipotezę H_0: wygenerowany rozkład jest rozkładem $T(\mu, \Delta)$ wobec H_1 że nie jest to prawdą. Korzystając z odpowiednich tabel statystycznych proszę sprawdzić czy nasza hipoteza jest prawdziwa na poziomie istotności $\alpha = 0.05$ (α jest prawdopodobieństwem pierwszego rodzaju czyli prawdopodobieństwem odrzucenia hipotezy H_0 gdy ta jest prawdziwa). W tym celu definiujemy obszar krytyczny testu:

$$\Phi = \{\boldsymbol{X} : \chi^2(\boldsymbol{X}) > \varepsilon\} \tag{7}$$

   gdzie: $\boldsymbol{X} = \{x_1, x_2, \ldots, x_n\}$ jest ciągiem liczb pseudolosowych, $\chi^2(\boldsymbol{X})$ wartością statystyki dla danego ciągu $\boldsymbol{X}$, ε jest poziomem krytycznym danego rozkładu dla określonej liczby stopni swobody i założonego poziomu istotności (należy odczytać z tabel statystycznych). Liczbę stopni swobody określamy jako $\nu = K - r - 1$, gdzie: K jest liczbą podprzedziałów, a $r = 2$ jest liczbą parametrów testowanego rozkładu (μ i Δ). Jeśli $\chi^2 < \varepsilon$ to stwierdzamy że dla zadanego poziomu istotności nie ma podstaw do odrzucenia hipotezy H_0.