Podstawowe typy rozkładów jednowymiarowych zmiennych losowych

Rozkład jednostajny (prostokątny): $X \sim U(a, b)$

- parametry: $a, b \in \mathbb{R} \wedge a < b$
- funkcja gęstości:

$$f(x) = I_{[a,b]}(x) \frac{1}{b-a}$$

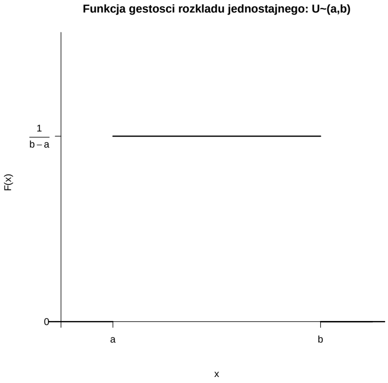
przy czym I jest funkcją charakterystyczną

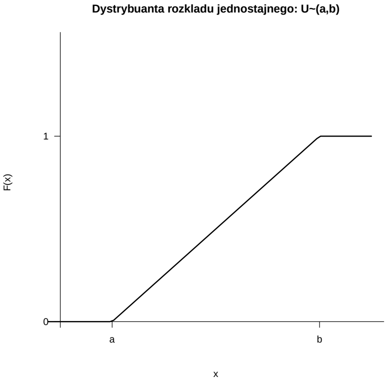
$$I_{[a,b]}(x) = \begin{cases} 0 & \text{gdy } x \notin [a,b] \\ 1 & \text{gdy } x \in [a,b] \end{cases}$$

- dystrybuanta:

$$F(x) = I_{[a,b)}(x) \frac{x-a}{b-a} + I_{[b,+\infty)}(x)$$

- wykresy funkcji gęstości i dystrybuanty rozkładu jednostajnego:





Rozkład normalny $X \sim N(\mu, \sigma)$

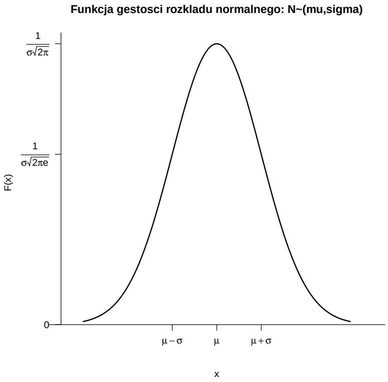
- parametry: $\mu \in \mathbb{R}$ - wartość oczekiwana, $\sigma \in \mathbb{R}_+$ - odchylenie standardowe
- funkcja gęstości:

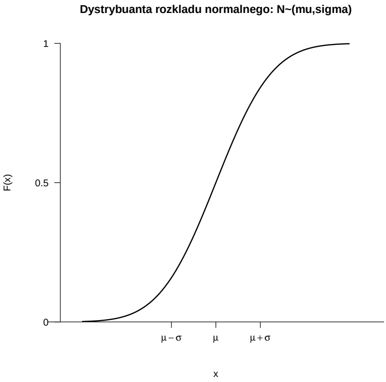
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

- dystrybuanta:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} dt$$

- wykresy funkcji gęstości i dystrybuanty rozkładu normalnego:





Generowanie liczb pseudolosowych dla rozkładu jednostajnego z parametrami (0,1) – *Linear Congruential Generator* (LCG)

Niech $a, b, p \in \mathbb{N}_0$ - ustalone parametry oraz $x_0 \in \mathbb{N}_0$ - ustalona wartość początkowa ciągu nazywana ziarnem (*seed*)

Dla odpowiednio dobranych parametrów a, b, p, x_0, wykorzystując wartości rekurencyjnie zadanego ciągu $x_{i+1} = (ax_i + b) \mod p$, można wyznaczyć liczby $u_i = \frac{x_i}{p}$, mające zbliżone własności do tych wylosowanych z rozkładu jednostajnego $U(0, 1)$.

Przy czym $\mod$ jest funkcją *modulo* przyporządkowującą resztę z dzielenia liczby $ax_i + b$ przez p, np. $15 \mod 4 = 3$.

Liczby pseudolosowe są generowane w sposób deterministyczny, a nie losowy (stąd też ich określenie). Tym samym ustalając takie same wartości parametrów generatora uzyskuje się za każdym razem identyczne wartości pseudolosowe.
Przyjmuje się, iż w celu zredukowania zjawiska cykliczności parametr p powinien być bardzo dużą liczbą, tak samo jak a, które powinno być względnie liczbą pierwszą z p. Wybór wartości parametru b nie jest tak istotny i często przyjmuje się dla niego wartość 0.

Przykładowy wybór parametrów generatora liczb pseudolosowych:
$a = 427419669081$, $p = 999999999989$, $b = 0$.
Natomiast zazwyczaj parametrem generatora liczb pseudolosowych modyfikowanym przez analityka jest wartość ziarna x_0, przy pozostałych parametrach niezmienionych :

- w celu uzyskania w kolejnych próbach identycznych ciągów liczb pseudolosowych, zakłada się w każdej z prób taką samą wartość ziarna,
- jeżeli w kolejnych próbach chce się uzyskać odmienne postaci ciągów liczb pseudolosowych, można w kolejnych próbach przypisać wartość ziarna równą ilości czasu zużytego przez procesor od rozpoczęcia aktualnej sesji.

Generowanie liczb pseudolosowych dla innych rozkładów jednowymiarowych – *inversion sampling*

Dystrybuanta F jest funkcją opisującą prawdopdobieństwo zdarzenia, że dana zmienna losowa X, o określonym rozkładzie, przyjmie wartości mniejsze od $x \in \mathcal{X}$: $F(x) = P\{X \leqslant x\}$.
Funkcja dystryburanty F może przyjmować wartości z przedziału $[0, 1]$, tzn. $F : \mathcal{X} \to [0, 1]$.
Funkcja odwrotna do dystrybuanty $F^{-1} : [0, 1] \to \mathcal{X}$ umożliwia odczytanie wartości zmiennej, dla której dystrybuanta przyjmuje wartość z przedziału $[0, 1]$, opisującą prawdopodobieństwo.

Tak więc do generowania liczb pseudolosowych dla rozkladów innych niz jednostajny można, wykorzystując definicję funkcji odwrotnej do dystrybuanty zakładanego rozkładu, skorzystać z generatora liczb pseudolosowych z rozkładu $U(0, 1)$: $x_i = F^{-1}(u_i)$, gdzie u_i - liczba pseudolosowa z generatora dla $U(0, 1)$, x_i - liczba pseudolosowa z zakładanego rozkładu (np. normalnego).