# Cvičení 10: Generování realizací spojité náhodné veličiny

## Metoda inverzní transformace

Pro generování realizací některých spojitých náhodných veličin se používá metoda inverzní transformace, která je založena na dvou následujících větách.

**Věta 1.:** Nechť spojitá náhodná veličina X má rostoucí distribuční funkci $\Phi(x)$ (tzn., že k ní existuje inverzní funkce $\Phi^{-1}(y)$ – tzv. kvantilová funkce). Pak transformovaná náhodná veličina $Y = \Phi(X)$ má rozložení Rs(0,1).

**Důkaz:** Označme $\Phi_*(y)$ distribuční funkci náhodné veličiny Y.

$\Phi_*(y) = P(Y \leq y) = P(\Phi(X) \leq y) = P\left(X \leq \Phi^{-1}(y)\right) = \Phi\left(\Phi^{-1}(y)\right) = y$ pro $y \in (0,1)$,

$\Phi_*(y) = 0$ pro $y \in (-\infty, 0\rangle$, $\Phi_*(y) = 1$ pro $y \in \langle 1, \infty)$

Tedy Y ~ Rs(0,1).

**Věta 2.:** Nechť X ~ Rs(0,1) a nechť $\Phi$ je rostoucí spojitá distribuční funkce. Pak transformovaná náhodná veličina $Y = \Phi^{-1}(X)$ má distribuční funkci $\Phi$.

**Důkaz:** Označme $\Phi_*(y)$ distribuční funkci náhodné veličiny Y.

$\Phi_*(y) = P(Y \leq y) = P\left(\Phi^{-1}(X) \leq y\right) = P(X \leq \Phi(y)) = \Phi(y)$, protože X ~ Rs(0,1).

**Postup:**
1. krok: Z rozložení Rs(0,1) vygenerujeme číslo u.
2. krok: Číslo u transformujeme vztahem $x = \Phi^{-1}(u)$.
Kroky 1 a 2 se opakují n-krát.

**Úkol 1.:** Vygenerujte 1000 realizací náhodné veličiny z rozložení Rs(-2, 5), nakreslete histogram a vypočtěte průměr a směrodatnou odchylku.

**Návod:** Hustota: $\varphi(x) = \begin{cases} \frac{1}{7} \text{ pro } x \in \langle -2, 5\rangle \\ 0 \text{ pro jinak} \end{cases}$, distribuční funkce: $\Phi(x) = \begin{cases} 0 \text{pro} x < -2 \\ \frac{x+2}{7} \text{pro} x \in \langle -2, 7\rangle \\ 1 \text{ pro } x > 7 \end{cases}$,

inverzní funkce k distribuční funkci (tzv.kvantilová funkce):

$\Phi^{-1}(u) = -2 + 7u$, kde $0 < u < 1$.

Vytvoříme nový datový soubor o dvou proměnných (nazveme je u a x) a 1000 případech. Do proměnné u uložíme náhodná čísla z intervalu (0, 1), tj. do Dlouhého jména této proměnné napíšeme =rnd(1). Proměnná x vznikne transformací proměnné u. Do Dlouhého jména proměnné x napíšeme =-2+7*u.
Vykreslíme histogram proměnné x. Histogram by měl sestávat ze sedmi téměř stejně vysokých obdélníků na intervalu $\langle -2, 5\rangle$.
Průměr a směrodatnou odchylku proměnné x vypočteme pomocí cesty Statistiky – Základní statistiky a tabulky – Popisné statistiky. Průměr by měl být blízký 1,5, směrodatná odchylka 2,02 (proč?).

**Úkol 2.:** Vygenerujte 1000 realizací náhodné veličiny z rozložení Ex(2) nakreslete histogram a vypočtěte průměr a směrodatnou odchylku.

**Návod:** Hustota: $\varphi(x) = \begin{cases} \lambda e^{-\lambda x} \text{ pro } x > 0 \\ 0 \text{ pro } x \le 0 \end{cases}$,

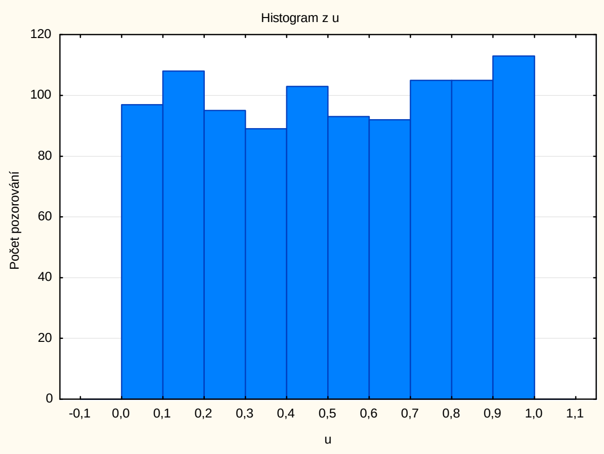
distribuční funkce: $\Phi(x) = \begin{cases} 1 - e^{-\lambda x} \text{ pro } x > 0 \\ 0 \text{ pro } x \le 0 \end{cases}$,

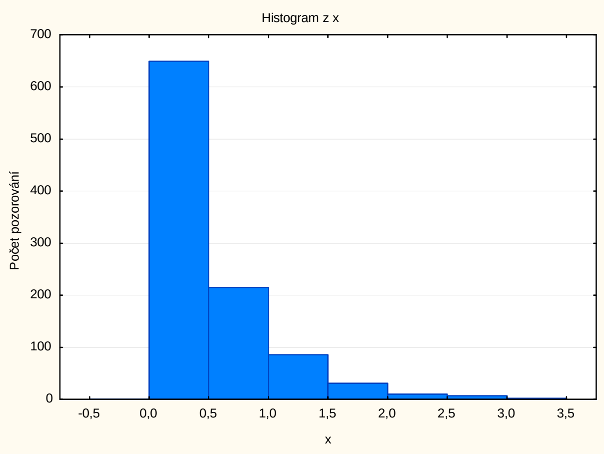
inverzní funkce k distribuční funkci (tzv.kvantilová funkce):

$\Phi^{-1}(u) = -\frac{1}{\lambda}\ln(1-u)$, kde $0 < u < 1$.

Vytvoříme nový datový soubor o dvou proměnných (nazveme je u a x) a 1000 případech. Do proměnné u uložíme náhodná čísla z intervalu (0, 1), tj. do Dlouhého jména této proměnné napíšeme =rnd(1). Proměnná x vznikne transformací proměnné u. Do Dlouhého jména proměnné x napíšeme =-0,5*log(1-u).
Vykreslíme histogramy proměnných u a x. První histogram by měl sestávat z 10 téměř stejně vysokých obdélníků na intervalu $\langle 0,1 \rangle$, tvar druhého histogramu by se měl blížit exponenciále.
Průměr proměnné x by se měl blížit 0,5, směrodatná odchylka rovněž (proč?)

**Upozornění:** V systému STATISTICA lze realizace náhodné veličiny s exponenciálním rozložením generovat též pomocí funkce RndExp(lambda).



**Úkol k samostatnému řešení:** Vygenerujte 1000 realizací náhodné veličiny s hustotou

$$\varphi(x) = \begin{cases} 4x^3 \text{ pro } 0 \le x \le 1 \\ 0 \text{ jinak} \end{cases}$$ a nakreslete histogram.

**Boxova – Müllerova transformace pro generování realizací náhodné veličiny s normálním rozložením:**

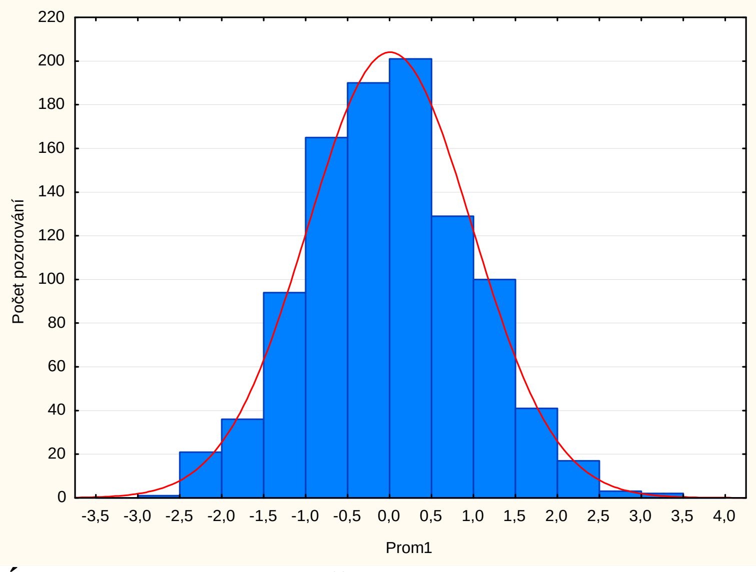
Nechť $x_1$, $x_2$ jsou dvě nezávislá čísla z Rs(0,1). Pak transformovaná čísla

$$z_1 = \sqrt{-2\ln x_1}\cos 2\pi x_2,\ z_2 = \sqrt{-2\ln x_1}\sin 2\pi x_2$$

lze považovat za realizace náhodné veličiny s rozložením N(0,1).

**Úkol 3.:** Pomocí B-M transformace vygenerujte 1000 realizací náhodné veličiny se standardizovaným normálním rozložením a nakreslete histogram.

**Návod:** Vytvoříme nový datový soubor o 4 proměnných (nazveme je x1, x2, z1, z2) a 500 případech. Do Dlouhého jména proměnných x1 a x2 napíšeme =rnd(1), do Dlouhého jména proměnné z1 napíšeme =sqrt(-2*log(x1))*cos(2*pi*x2) a do Dlouhého jména z2 napíšeme =sqrt(-2*log(x1))*sin(2*pi*x2). Dále vytvoříme nový datový soubor o jedné proměnné (nazveme ji z) a 1000 případech. Do případů 1 – 500 okopírujeme hodnoty proměnné z1 a do případů 501 – 1000 okopírujeme hodnoty proměnné z2. Vykreslíme histogram proměnné z s proloženou Gaussovou křivkou.




**Úkol k samostatnému řešení:** Proměnnou z vytvořenou v úkolu 2 transformujte na proměnnou s rozložením N(1, 1,5)

Upozornění: V systému STATISTICA lze realizace náhodné veličiny s normálním rozložením N(0, $\sigma^2$) generovat pomocí funkce RndNormal($\sigma$).