



# Modelování a simulace 1

## 8. lekce - Generování pseudonáhodných čísel

**Michal Janošek**

Department of Informatics and Computers

Faculty of Science

University of Ostrava

Ostrava, Czech Republic

`michal.janosek@osu.cz`

November 14, 2017



[http://i.telegraph.co.uk/multimedia/archive/01838/lottery\\_1838360b.jpg](http://i.telegraph.co.uk/multimedia/archive/01838/lottery_1838360b.jpg)

# Generování pseudonáhodných čísel

- Náhodná vs. pseudonáhodná čísla
- Kritéria náhodnosti
- Algoritmy pro generování pseudonáhodných čísel
- Generování pseudonáhodných čísel z daného rozdělení
- Implementace generátoru
- Testování generátoru

# Kritéria náhodnosti

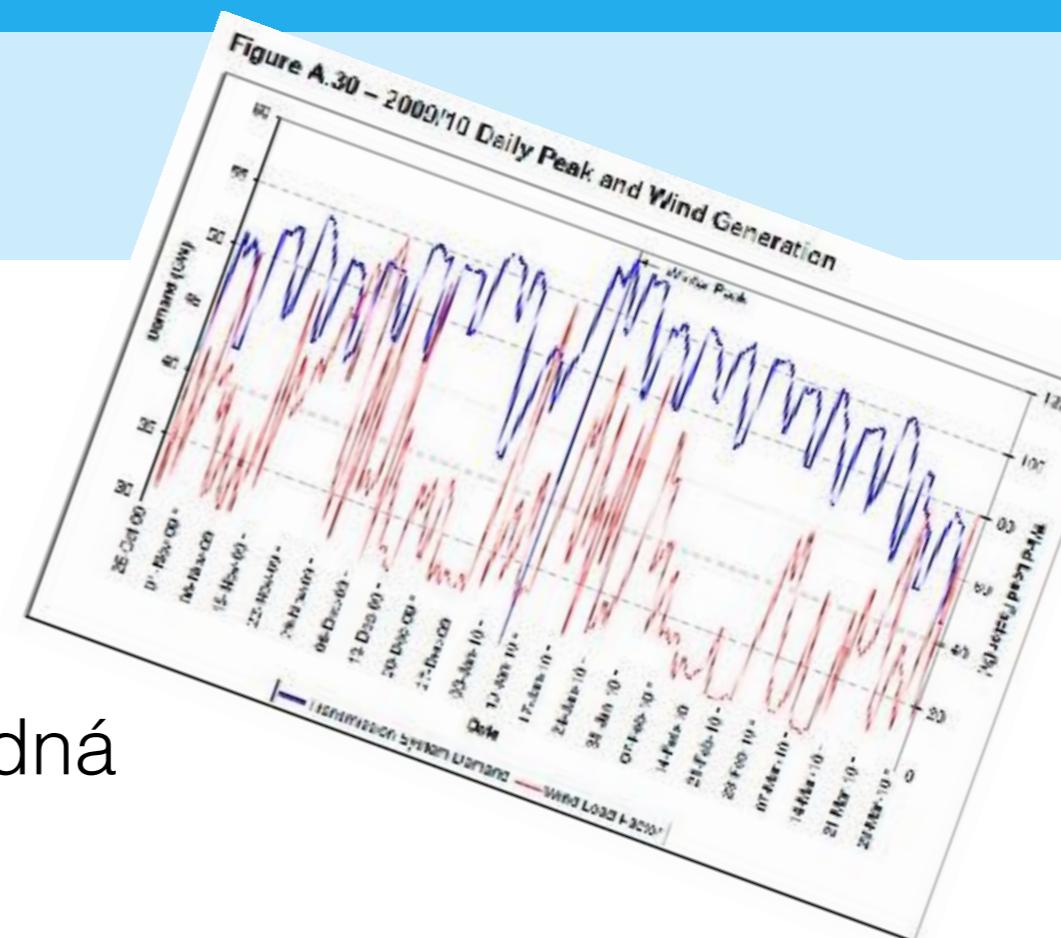
- procesy v reálném světě mají náhodný charakter
- analýza reálného systému → náhodný charakter některého z procesů → respektování při konstrukci simulačního modelu
- hledání algoritmů vytvářející posloupnosti náhodných čísel splňující **základní kritéria náhodnosti**
  - např. ve tvaru pravých (jmenovatel větší než čitatel) desetinných zlomků s pevným počtem s číslic za desetinnou čárkou
  - kde  $a_i \in \{0, 1, \dots, 9\}$ ,  $i=1, 2, \dots, s$ 
    - každá číslice vybraná z množiny  $\{0, 1, \dots, 9\}$  má stejnou pravděpodobnost být vybrána
    - výběr číslice  $a_i$  nemá vliv na výběr následující číslice  $a_{i+1}$

# Generování náhodných čísel pro simulační model

- při konstrukci generátoru vycházíme z experimentálních dat
  - využítí teoretické distribuční funkce (popisující pravděpodobnost rozdělení dat)
  - konstrukce empirické distribuční funkce na základě rozdělení četnosti uvažovaných dat

# Generátor náhodných čísel

- zařízení/procedura gen. náhodná čísla
  - jsou náhodná, připomínají čísla náhodná
  - postrádají jakýkoliv vzor
- vstupní parametry
  - poč. hodnota, min.-max. hodnota, max. rozpětí mezi generovanými čísly
- PRN generátor (pseudonáhodná čísla)
- TRN generátor (pravý generátor, fyz. jev)



[http://www.windbyte.co.uk/ims/windpower/ng\\_winter0910\\_wind\\_demand.jpg](http://www.windbyte.co.uk/ims/windpower/ng_winter0910_wind_demand.jpg)

# Generování (pseudo)náhodných čísel

- měření fyzikálního jevu (hardwareové metody)
  - předpoklad, že je náhodný
  - kompenzace odchylek
- výpočetní algoritmy (softwarové metody)
  - dlouhé sekvence zdánlivě náhodných čísel
  - vstupní hodnota: klíč
  - deterministicky vypočtena ze vstupní hodnoty
  - hodnota je předvídatelná
- rozpoznání pravých a pseudonáhodných čísel
  - metody statistické analýzy

# Fyzikální - hardwareové metody

- hod kostkou, mincí, ruleta
  - pomalé neúčinné
  - množství fyz. principů (nepředvídatelnost náh. chování)
  - atomárních a subatomárních jevů (kvantová mechanika)
  - radioaktivní látky (rozpad)
  - měření špiček v síti
  - měření šumů
  - vstupů od uživatele (pohyb myši, stiskem kláves...)

fyzikální generátor náh. čísel  
<http://www.random.org/>

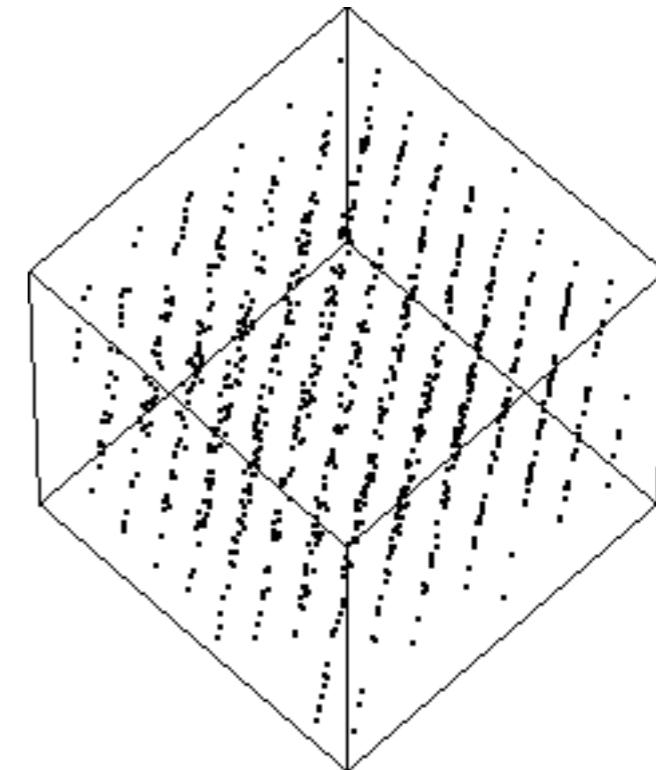


# Výpočetní - softwarové, rekurentní vztahy

- algoritmy vytvářející dlouhé řetězce
- zdánlivě náhodné rozdělení
- sekvence se opakují
- kvalita se snižuje

$$x_{n+1} = x_n^2 \bmod M,$$

$$X_{n+1} = (aX_n + b) \bmod m$$



( $a=65539$ ,  $b=0$ ,  $m=2^{31}$ ,  $x_0$  liché)  
generátor RANDU (60. léta)

- odlišné hodnoty násobícího koeficientu
- programy - speciální knihovny
  - hodiny v počítači
  - dostatečné pro běžné aplikace, nedostatečné pro šifrování

zdroj	$m$	$a$	$c$	výstup
Numerical Recipes	$2^{32}$	1 664 525	1 013 904 223	
Borland C/C++	$2^{32}$	22 685 477		1 bity 30–16 v rand(), 30–0 v frand()
gcc (GCC)	$2^{32}$	1 103 515 245		12 345 bity 30–0
ANSI C: Watcom, Digital Mars, CodeWarrior, IBM VisualAge C/C++	$2^{32}$	1 103 515 245		12 345 bity 30–16
Borland Delphi, Virtual Pascal	$2^{32}$	134 775 813		1 bity 63–32 za (seed * L)
Microsoft Visual/Quintic C/C++	$2^{32}$	214 013		2 531 011 bity 30–16
Apple CarbonLib (Park & Miller)	$2^{32} - 1$	16 807		0

[https://cs.wikipedia.org/wiki/Line%C3%A1rn%C3%AD\\_kongruentn%C3%AD\\_gener%C3%A1tor](https://cs.wikipedia.org/wiki/Line%C3%A1rn%C3%AD_kongruentn%C3%AD_gener%C3%A1tor)

# Generování pseudonáhodných čísel z daného rozdělení

- tyto generátory vytvářejí pseudonáhodné posloupnosti  $\{x_i\}$ 
  - členy pocházejí z rovnoměrného rodělení na intervalu  $\langle 0, 1 \rangle$
  - pomocí jednoduché lineární transformace se generuje pseudonáhodná posloupnost z rovnoměrného rozdělení na intervalu  $\langle A, B \rangle$

$$y_i = A + (B - A) x_i \quad (i = 1, 2, \dots)$$

# Implementace generátoru pseudonáhodných čísel v simulačním programu

- procedurou
  - provádí výpočet podle zvoleného algoritmu generátoru
- speciální proměnnou
  - hnízdo resp. násada generátoru
    - zde se uchovává rekurentně měněná hodnota pro další volání procedury generátoru
- potřebujeme-li simulovat více navzájem nezávislých náhodných procesů, použije více hnízd generátoru
  - tato hnízda musí být různě inicializována

# Ověření a testování

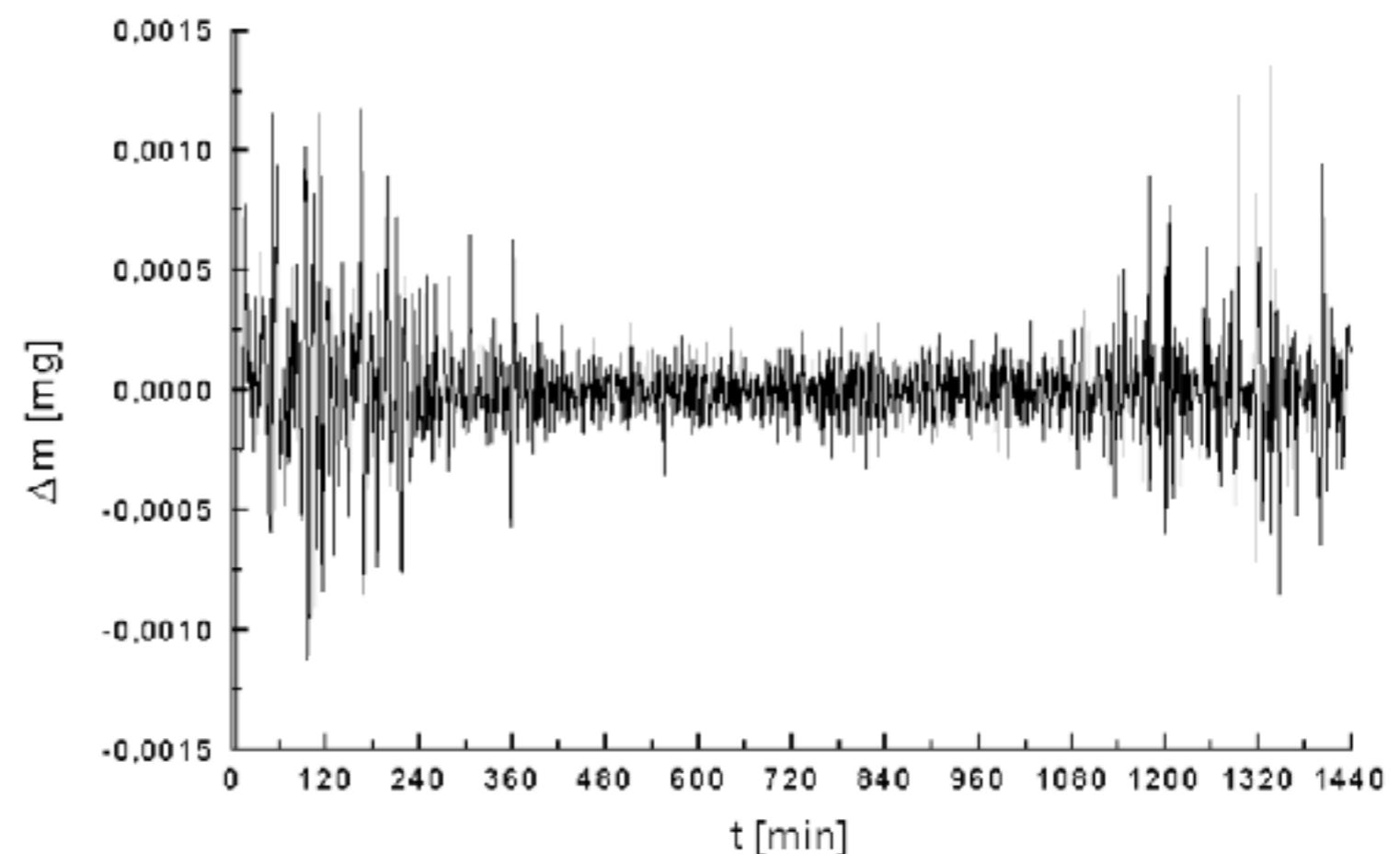
- ověřování i fyz. generovaných čísel
  - správná funkčnost (změna napětí, teploty, stáří zařízení)
  - softwarové chyby v kódu, hw defekty
  - testování generátoru
    - frekvenční test
    - testy autokorelace
    - sériový test
- ENT - program pro testování generátoru pseudonáhodných čísel  
<http://www.fourmilab.ch/random/>

# Periodicita generátoru

- inicializace generátoru - poč. stav (random seed)
- generuje vždy stejnou sekvenci čísel se stejným random seed
- perioda - maximum nad všemi počátečními stavy délky neopakující se sekvence
- periody pro některé výchozí stavy mohou být kratší než je očekáváno

# Aplikace

- důraz na proměnlivost systému
- šifrování
- poč. simulace
- bezp. systémy
- hazardní hry



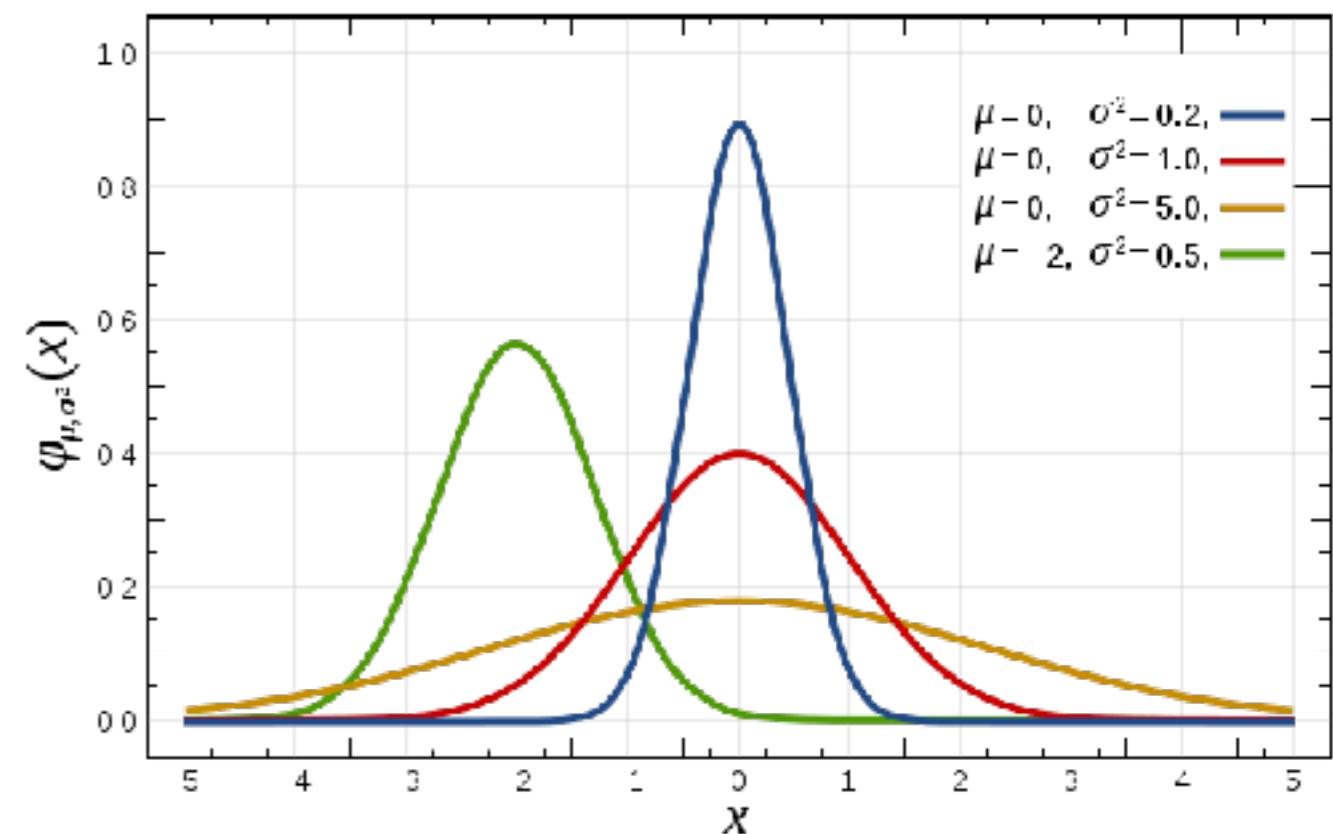
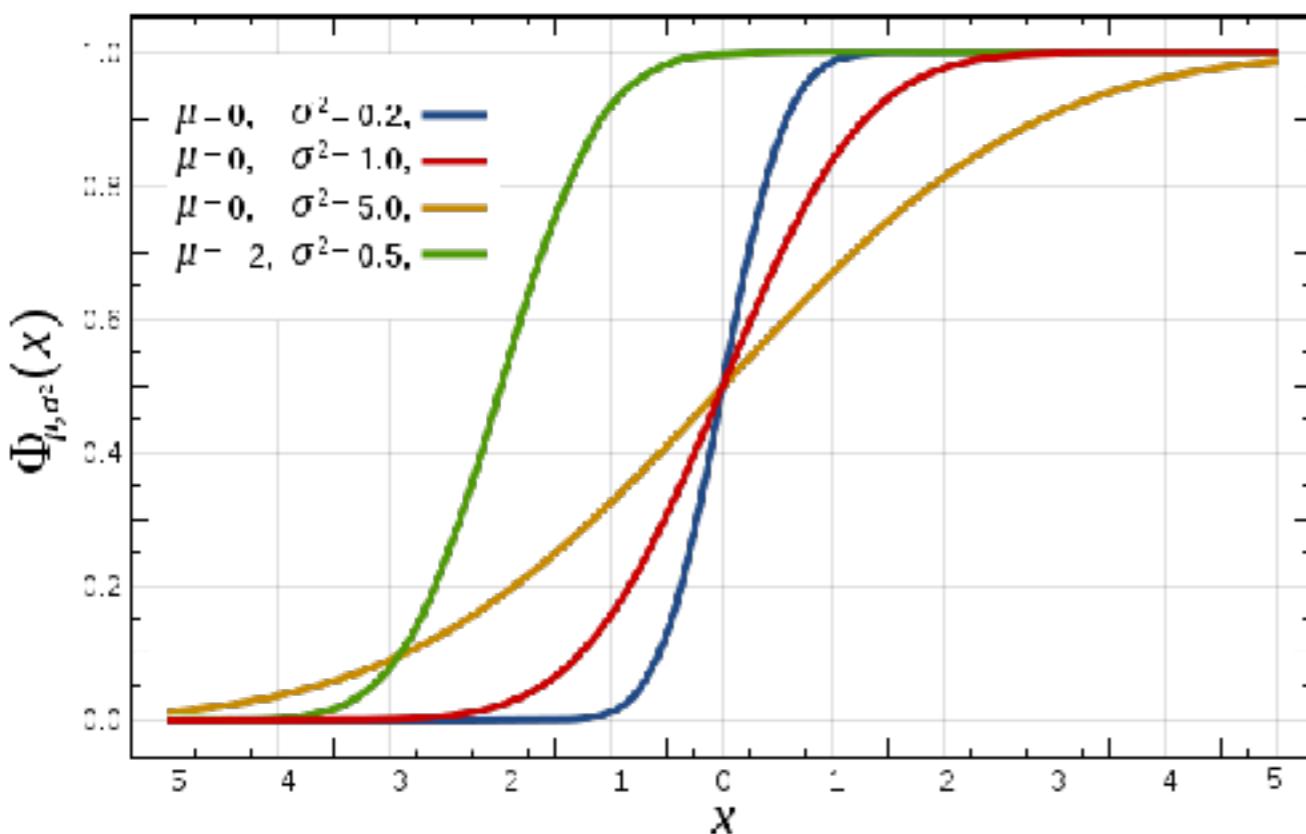
# Ja zvolit vhodnou distribuční funkci?

LAW, Averill M., 2016. *A tutorial on how to select simulation input probability distributions*

- všechny reálné systémy obsahují jeden nebo více zdrojů náhodnosti
- náhodný charakter vstupů simulačních programů
  - např. čekání ve frontě, zpracování výrobku
- simulační program generuje hodnoty dané specifikovanou distribuční funkcí
- Otázky:
  - Jak modelovat zdroj náhodnosti?
  - Jak zvolit vhodnou distribuční funkci, její parametry a ověřit vhodnost výběru této funkce?

# Ja zvolit vhodnou distribuční funkci?

[https://cs.wikipedia.org/wiki/Distribu%C4%8Dn%C3%AD\\_funkce](https://cs.wikipedia.org/wiki/Distribu%C4%8Dn%C3%AD_funkce)



- distribuční funkce vs. hustota pravděpodobnosti

# Ja zvolit vhodnou distribuční funkci?

LAW, Averill M., 2016. A tutorial on how to select simulation input probability distributions

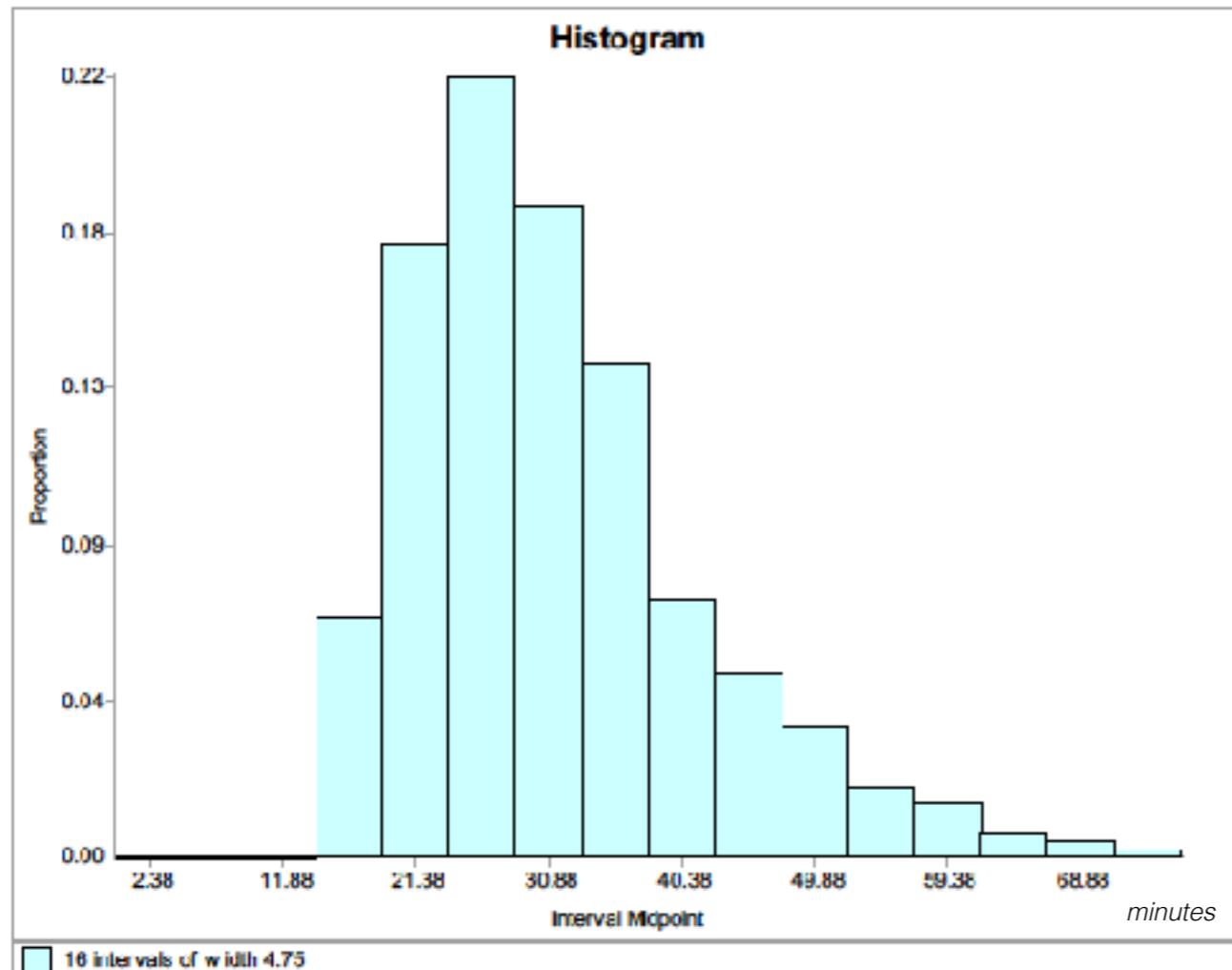


Figure 1: Histogram of 910 processing times for an automotive manufacturer.

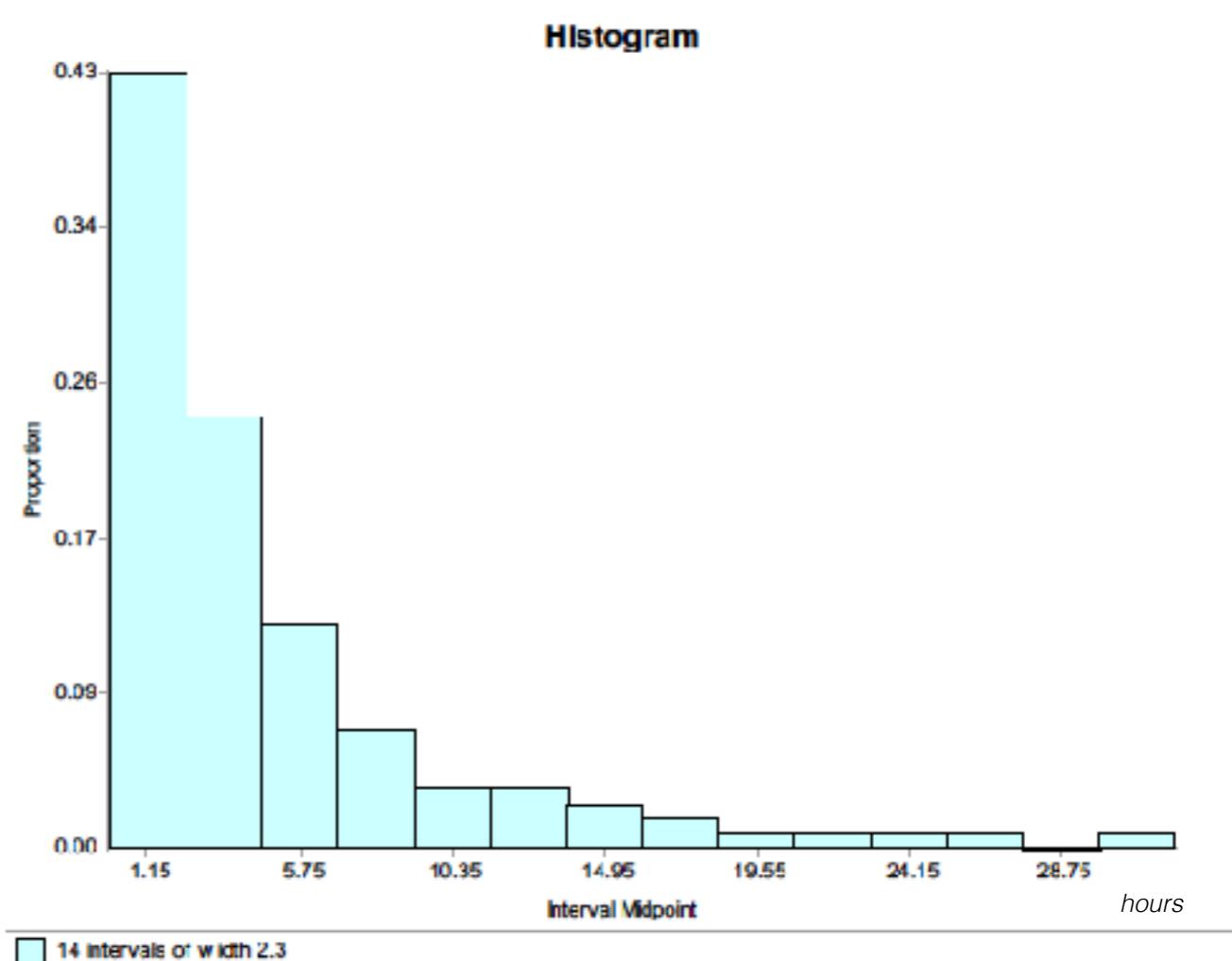
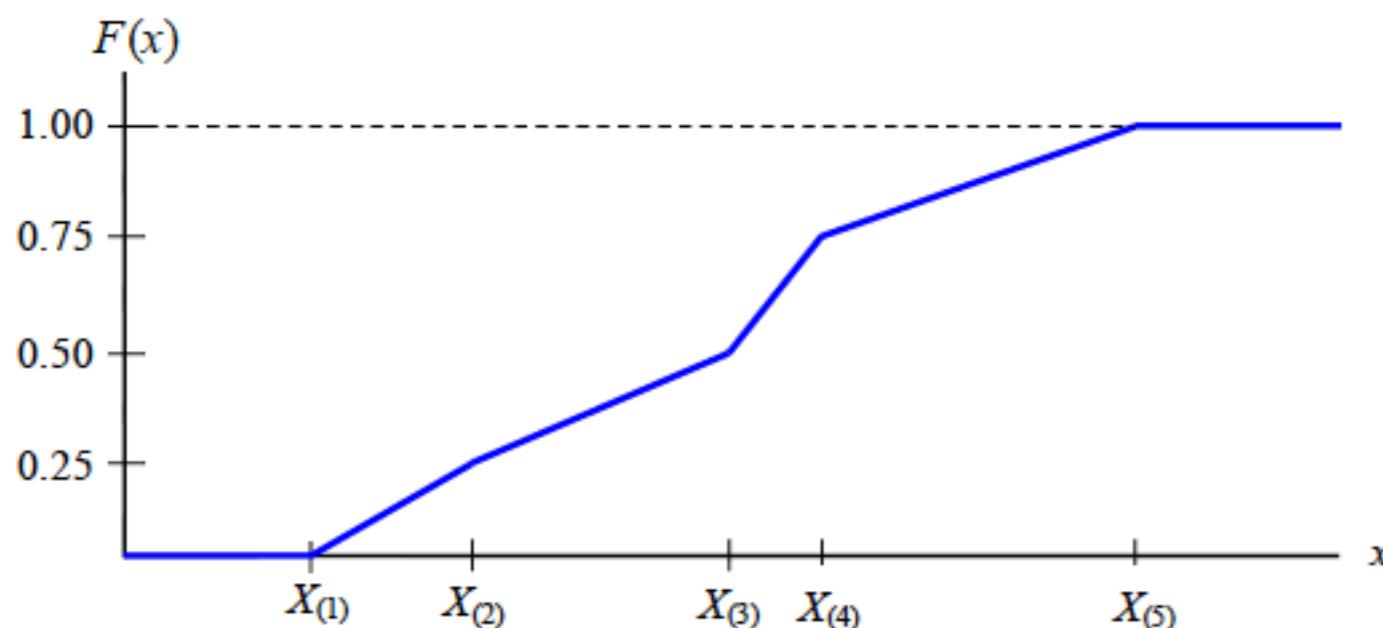


Figure 2. Histogram of 122 repair times for a U.S. Navy weapons system.

# Ja zvolit vhodnou distribuční funkci?



LAW, Averill M., 2016. A tutorial on how to select simulation input probability distributions

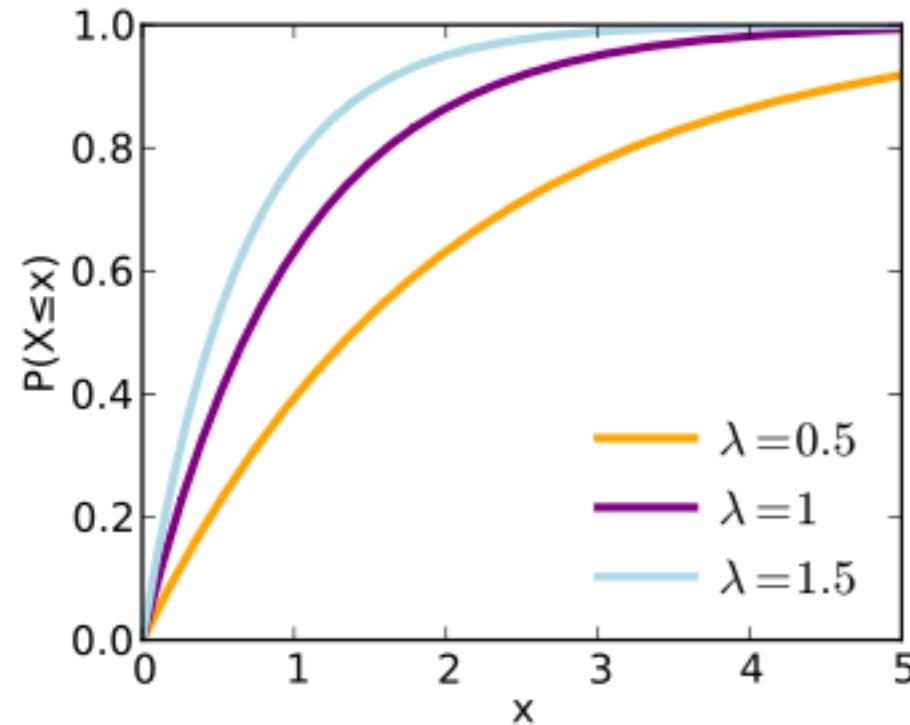
Figure 4: Continuous, piecewise-linear empirical distribution function

$$F(x) = \begin{cases} 0 & \text{if } x < X_{(1)} \\ \frac{i-1}{n-1} + \frac{x - X_{(i)}}{(X_{(i+1)} - X_{(i)})} & \text{if } X_{(i)} \leq x < X_{(i+1)} \text{ for } i = 1, 2, \dots, n-1 \\ 1 & \text{if } X_{(n)} \leq x \end{cases}$$

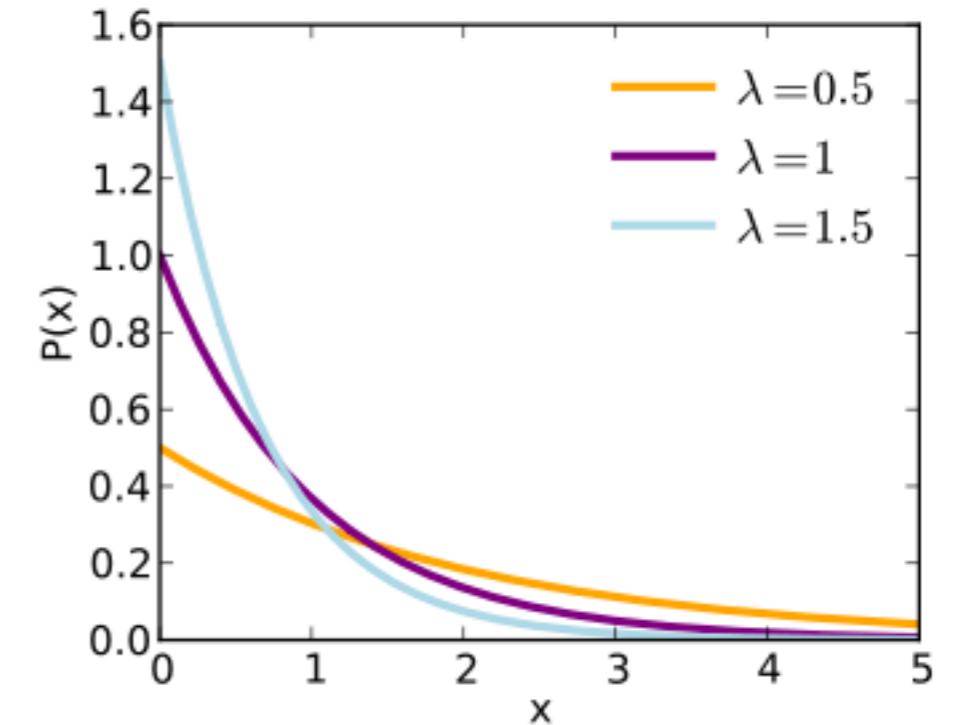
- empirická distribuční funkce
  - odhad skutečné distribuční funkce

# Ja zvolit vhodnou distribuční funkci?

[https://en.wikipedia.org/wiki/Exponential\\_distribution](https://en.wikipedia.org/wiki/Exponential_distribution)



$$F(x; \lambda) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0, \\ 0 & x < 0 \end{cases}$$



$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

- teoretická distribuční funkce
  - rodina distribučních funkcí
  - odhad parametrů distribuční funkce

# Ja zvolit vhodnou distribuční funkci?

LAW, Averill M., 2016. A tutorial on how to select simulation input probability distributions

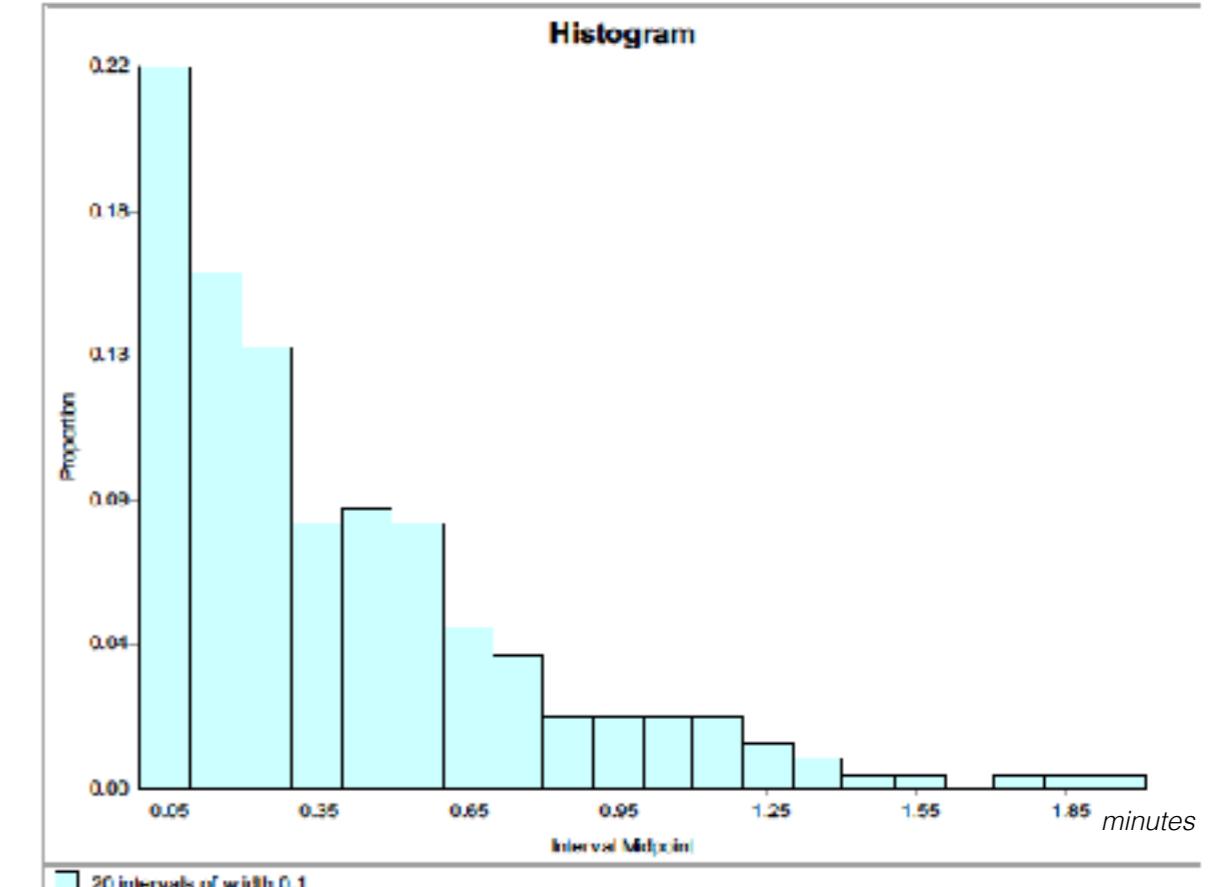
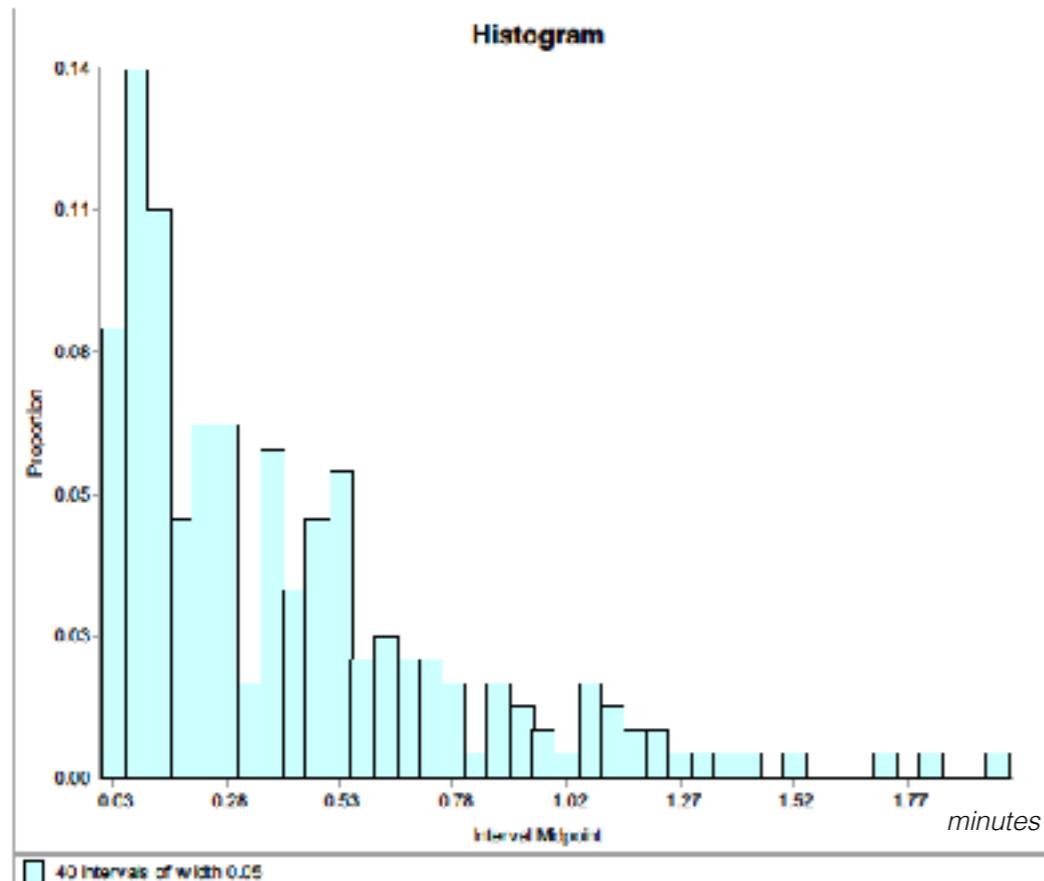


Figure 5: Histogram of 219 interarrival times to a drive-up bank with an interval width of 0.05.

Figure 3: Histogram of 219 interarrival times to a drive-up bank.

- histogram by měl být tzv. hladký

# Ja zvolit vhodnou distribuční funkci?

LAW, Averill M., 2016. A tutorial on how to select simulation input probability distributions

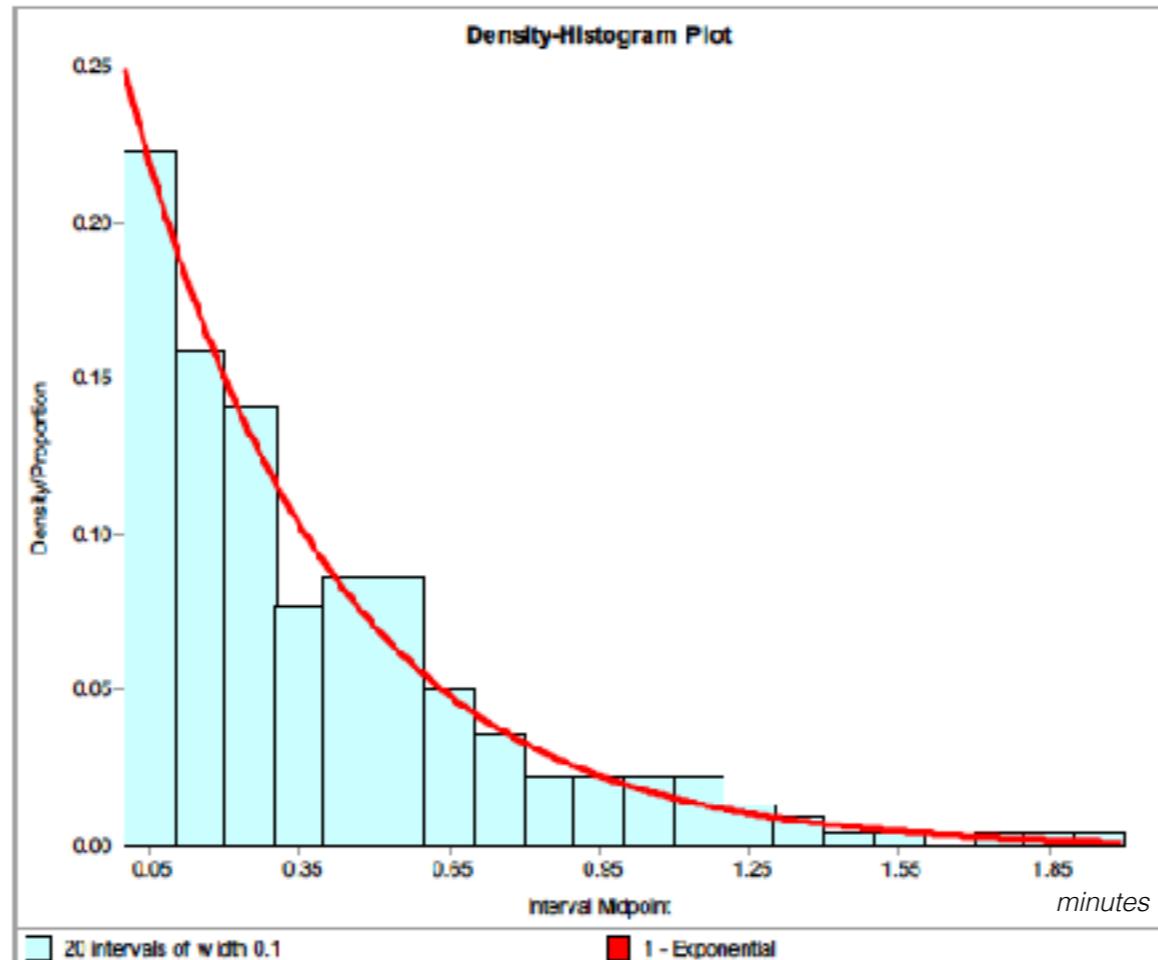


Figure 6: Density histogram plot for the fitted exponential distribution and the interarrival time data.

- jak reprezentativní je zvolená distribuční funkce?

# Použité obrázky

- [http://upload.wikimedia.org/wikipedia/commons/thumb/c/c7/Blue\\_lava\\_lamp.jpg/398px-Blue\\_lava\\_lamp.jpg](http://upload.wikimedia.org/wikipedia/commons/thumb/c/c7/Blue_lava_lamp.jpg/398px-Blue_lava_lamp.jpg)
- [http://en.wikipedia.org/wiki/File:Analyse\\_thermo\\_gravimetrique\\_bruit.png](http://en.wikipedia.org/wiki/File:Analyse_thermo_gravimetrique_bruit.png)