

Přednáška č. 3

SPOJITÁ NÁHODNÁ VELIČINA

4. 3. 2021

P. Pecherková

SPOJITÁ NÁHODNÁ VELIČINA

- veličina, která má nekonečný nebo nespočetný počet realizací

- Neúplný popis – popisná statistika

- Charakteristiky polohy
- Charakteristiky variability

- Úplný popis

- Hustota pravděpodobnosti (pdf)
- Distribuční funkce (cdf)

Příklad:

- Čas – nekonečné
- Teplota – nekonečné
- Cena – nespočetné
- Mzda – nespočetné
- Délka – nekonečné / nespočetné
- Hmotnost – nekonečné / nespočetné
- Vzdálenost – nekonečné
- Druhy vozidel – nespočetné
- Barvy - nespočetné

HUSTOTA PRAVDĚPODOBNOSTI

Def.: Pro spojitou náhodnou veličinu X definujeme hustotu pravděpodobnosti f_X jako reálnou funkci reálného argumentu vztahem

$$f_X(x) = \frac{dF_X(x)}{dx}$$

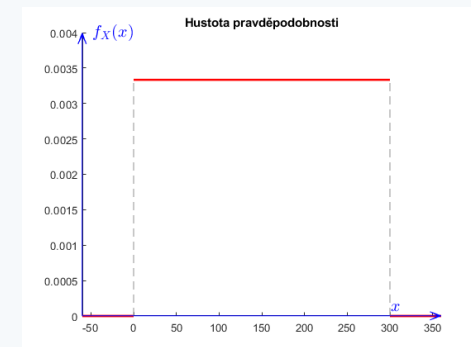
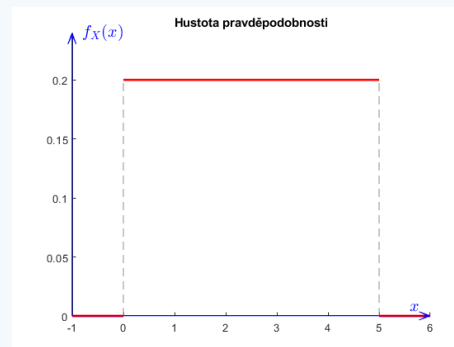
Hustota pravděpodobnostní funkce v bodě x_i se rovná 0 (udávaná pravděpodobnost je vztažena k intervalu).

Vlastnosti pravděpodobnostní funkce:

- Nezáporná na celém definičním oboru $f_X(x) \geq 0$
- její integrál přes celý definiční obor je jedna, tedy $\int_{-\infty}^{\infty} f_X(x) dx = 1$

Čekáme na tramvaj, která jezdí pravidelně každých 5 minut.
Pravděpodobnost, že přijde přesně za 1 min od mého příchodu je rovna 0 (může přijet za 1 min a 2 s, apod).

Pokud si to představíme v intervalech, tak pravděpodobnost, že přijede v 1. minutě je stejná jako v 2. minutě atd.



DISTRIBUČNÍ FUNKCE

Distribuční funkce F_X reálného argumentu x pro náhodnou veličinu X je definována vztahem

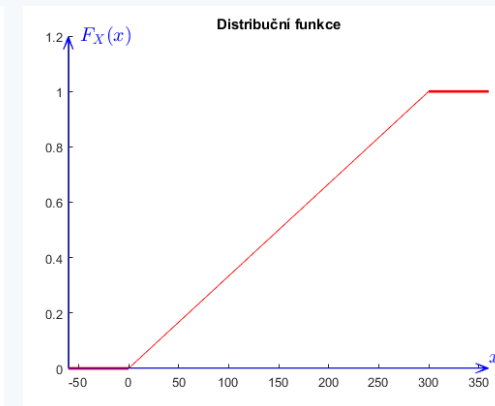
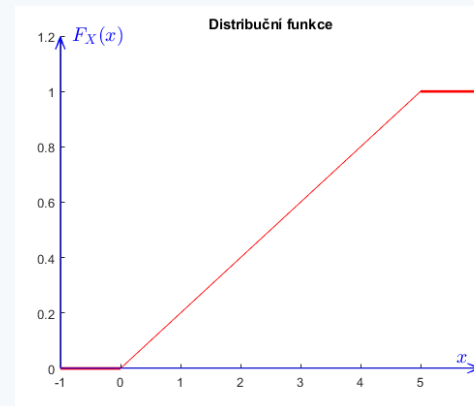
$$F_X(x) = P(X \leq x)$$

Distribuční funkce je součet pravděpodobností, které jsou menší nebo rovno danému bodu x .

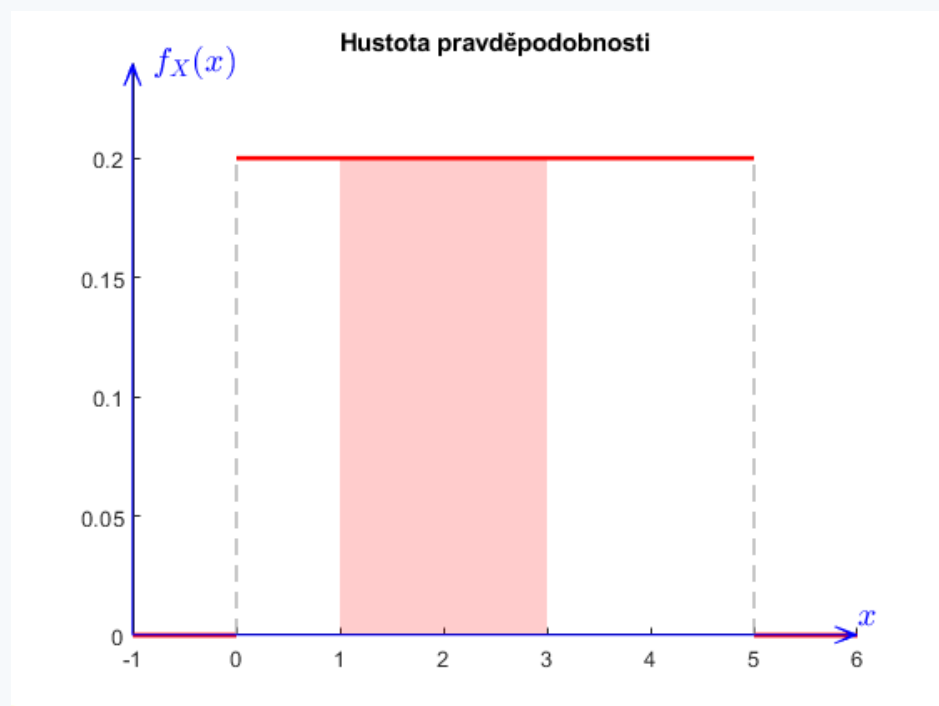
Vlastnosti distribuční funkce:

- Neklesající v celém oboru
- $\lim_{x \rightarrow -\infty} F(x) = 0$
- $\lim_{x \rightarrow \infty} F(x) = 1$
- $P(x \in (a, b)) = F(b) - F(a)$

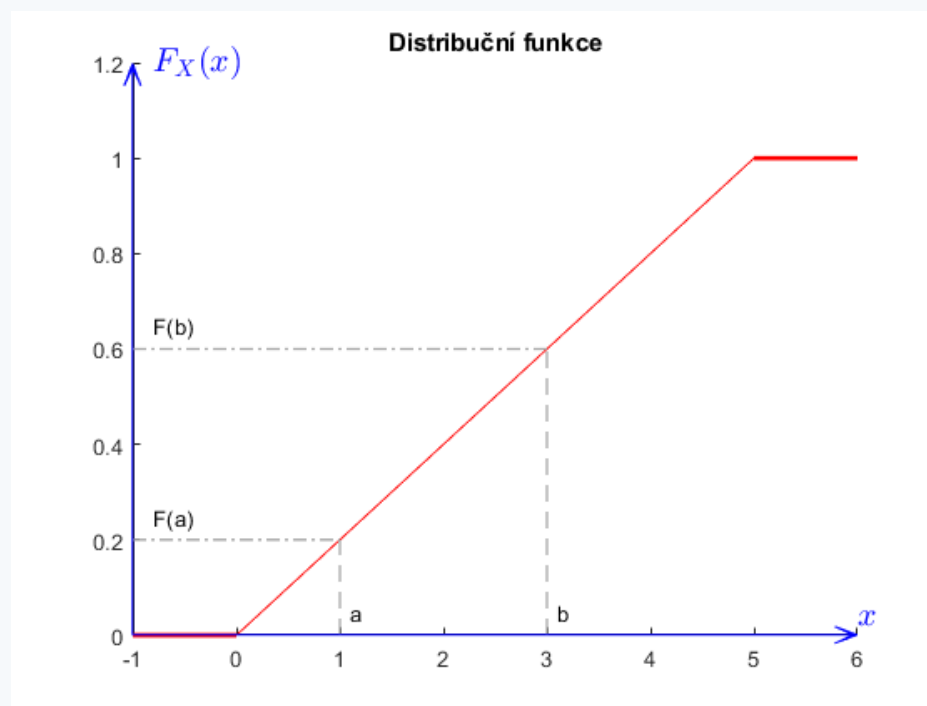
Čekáme na tramvaj, která jezdí pravidelně každých 5 minut.
Pravděpodobnost, že přijde přesně za 1 min od mého příchodu je rovna 0 (může přijet za 1 min a 2 s, apod).



S jakou pravděpodobností budu na metro čekat od 1 minuty do 3 minut?



$$P(x \in (a, b)) = \int_a^b f(x) dx = \int_1^3 f(x) dx = \frac{2}{5}$$



$$P(x \in (a, b)) = F(b) - F(a) = \frac{3}{5} - \frac{1}{5} = \frac{2}{5}$$

CHARAKTERISTIKY POLOHY

- Střední hodnota

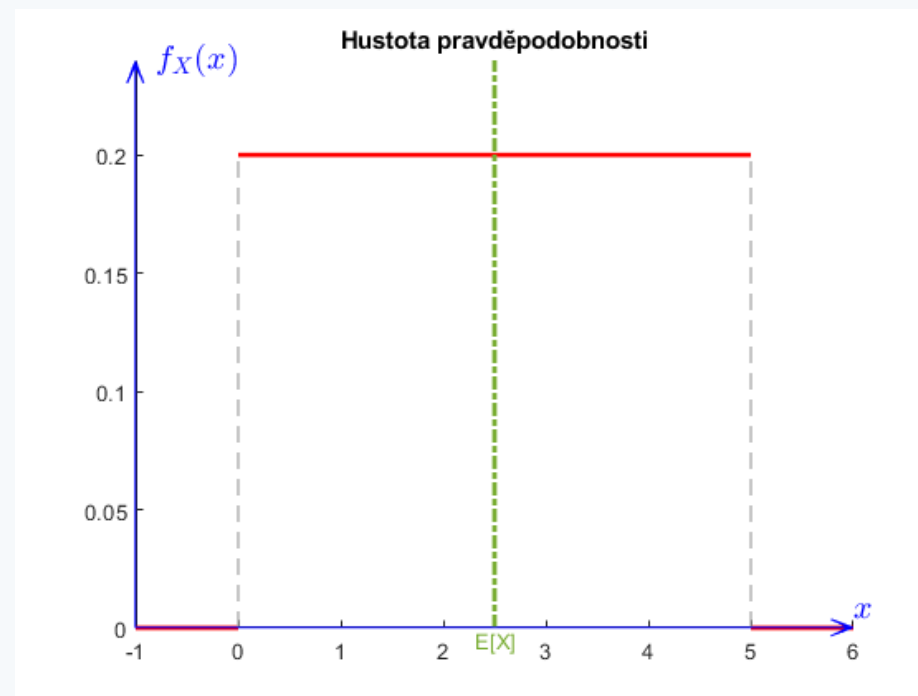
- $E[X] = \int_{-\infty}^{\infty} xf(x) dx$

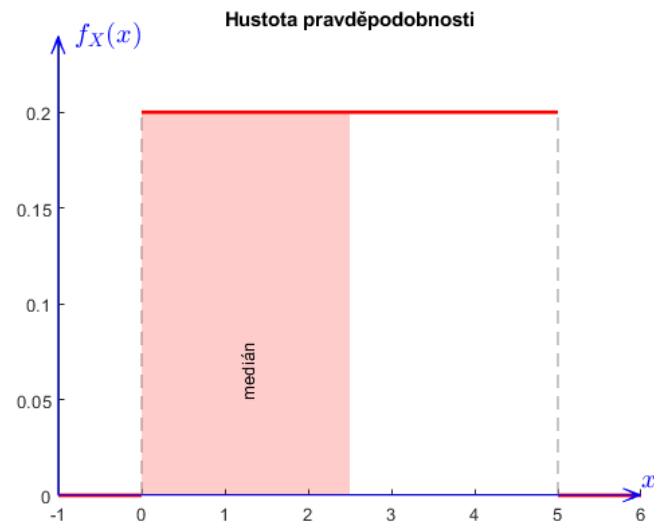
- $E[X] = \int_0^5 x \frac{1}{5} dx = \frac{1}{5} \left[\frac{x^2}{2} \right]_0^5 = \frac{1}{5} (12,5 + 0) = 2,5$

- Modus

- $\hat{x} = \arg \max(f(x))$

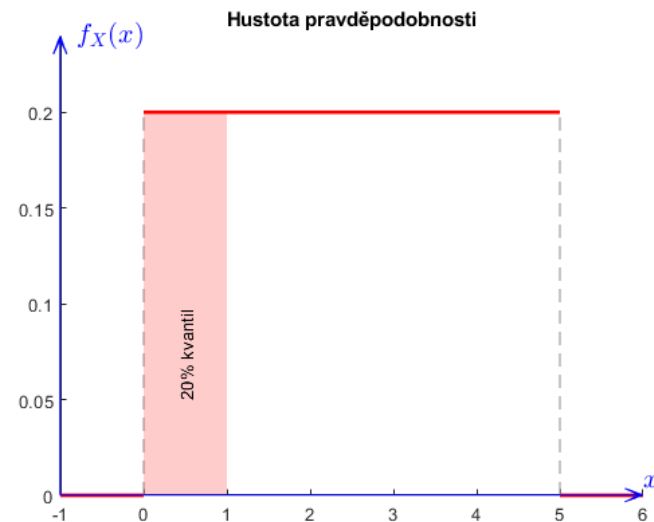
- $\hat{x}$ není (hledáme maximum funkce)





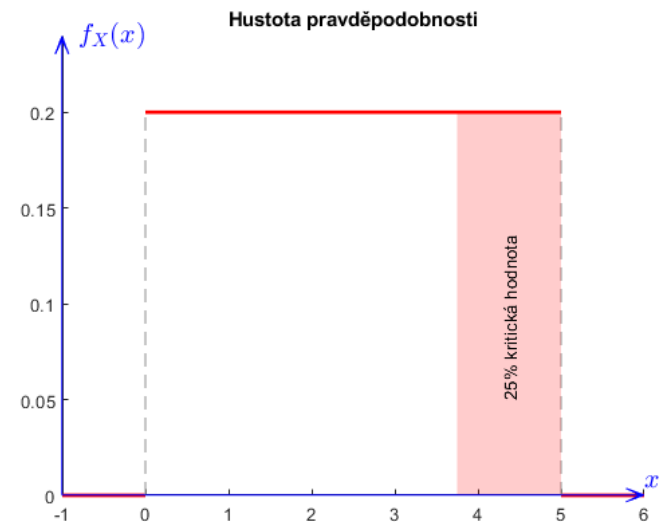
- Medián

- $\int_{-\infty}^{\tilde{x}_{0,5}} f(x) dx = 0,5$



- Kvantil

- $\int_{-\infty}^{\xi_{\alpha}} f(x) dx = \alpha$



- Kritická hodnota

- $\int_{z_{\alpha}}^{\infty} f(x) dx = \alpha$

CHARAKTERISTIKY VARIABILITY

- Rozptyl

- $D[x] = \int_{-\infty}^{\infty} (x - E[X])^2 f(x) dx$

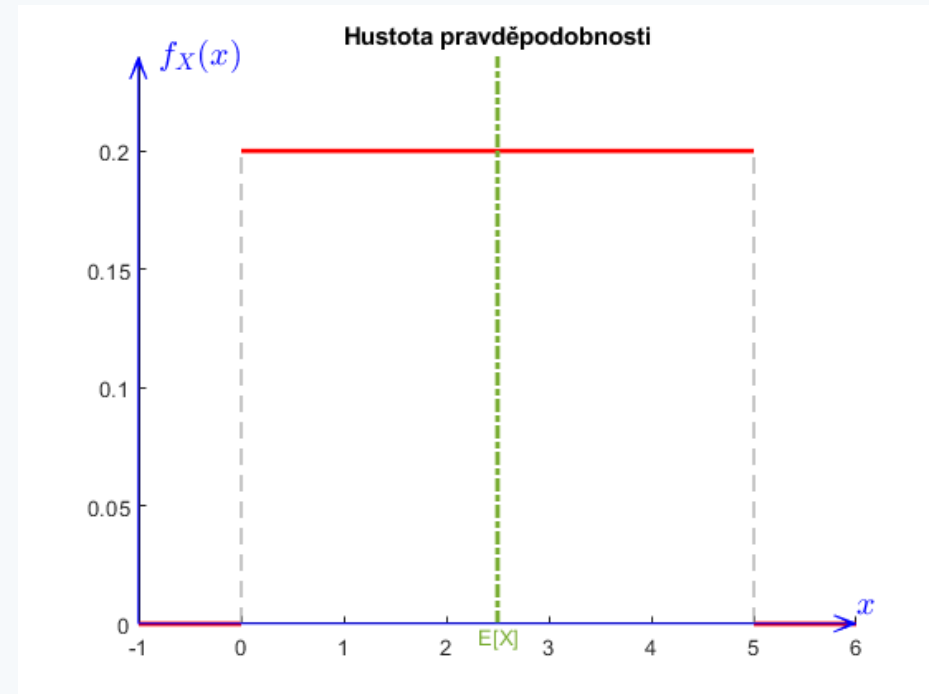
- $D[x] = \int_{-\infty}^{\infty} x^2 f(x) dx - E[X]^2$

- $D[x] = \int_0^5 x^2 f(x) dx - 2,5^2 = \frac{25}{3} - 6,25 = 2 \frac{1}{12}$

- Směrodatná odchylka

- $\sigma = \sqrt{D[X]}$

- $\sigma = \sqrt{2 \frac{1}{12}} = 1,44$

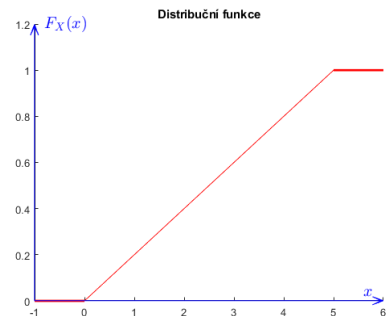


ROZDĚLENÍ SPOJITÉ NÁHODNÉ VELIČINY

- Některé chování náhodné veličiny je popsáno pomocí různých již známých rozdělení



Rovnoměrné rozdělení = $U(a, b)$

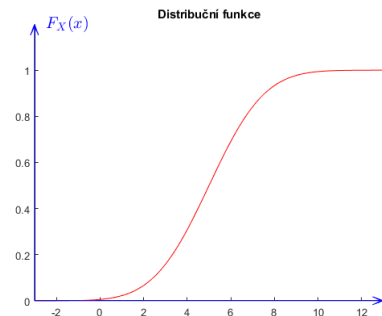
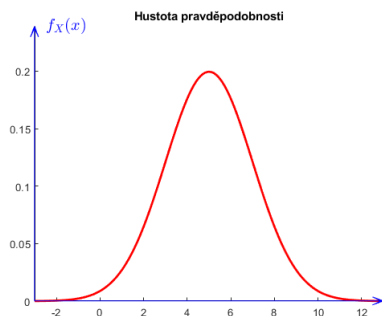


- Rovnoměrné rozdělení má dvě důležité charakteristik
 - všechny realizace jsou rovnocenné (žádný výsledek není preferován)
 - ostré hranice pro definiční obor (hranice nelze překročit)

- Hustota pravděpodobnosti:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{pro } x \in (a, b) \\ 0 & \text{jinde} \end{cases}$$
- Střední hodnota: $E[X] = \frac{a+b}{2}$
- Rozptyl: $D[X] = \frac{1}{12}(b-a)^2$

Normální rozdělení = $N(\mu, \sigma^2)$

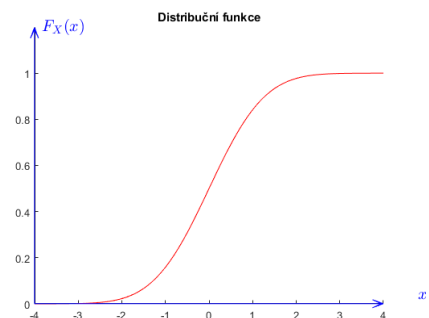
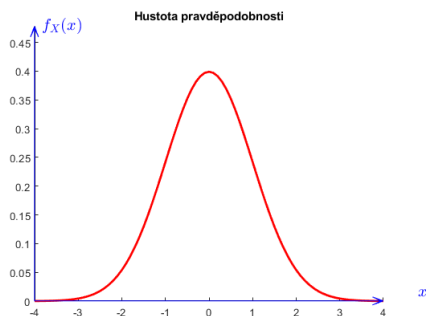


- Normální rozdělení je nejčastěji používaným rozdělením.
 - Dají se pro něj analyticky dopočítat i složité statistické úlohy
 - Existuje na nekonečném prostoru
- Hustota pravděpodobnosti:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad x \in \mathbb{R}$$

- Střední hodnota: $E[X] = \mu$
- Rozptyl: $D[X] = \sigma^2$

Normované normální rozdělení $N(0, 1)$

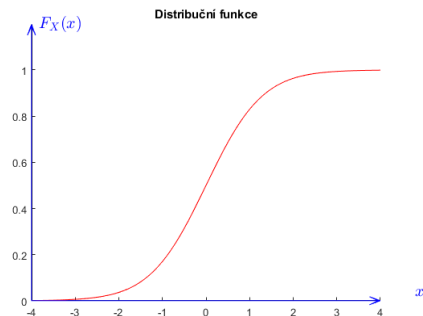
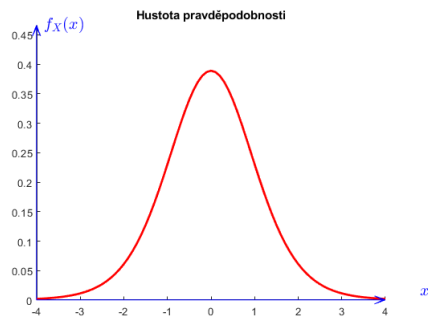


- Jedná se o speciální typ normálního rozdělení, kdy střední hodnota je rovna 0 a rozptyl se rovná hodnotě 1.
- Hustota pravděpodobnosti:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad x \in \mathbb{R}$$

- Střední hodnota: $E[X] = 0$
- Rozptyl: $D[X] = 1$

Studentovo t-rozdělení = $t(n)$

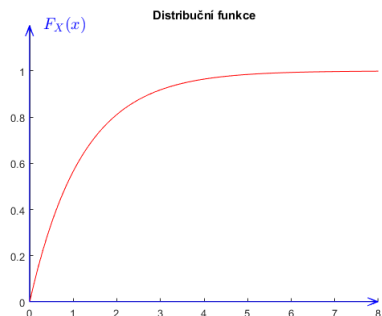


- Studentovo má velmi podobný tvar jako normální. Používá se v případě, kdy očekáváme chování jako u normálního rozdělení, ale neznáme rozptyl.
- n = počet stupňů volnosti ... počet dat - 1
- Je to umělé rozdělení a počítá se jako podíl normovaného normálního a χ^2 (chí-kvadrát) rozdělení
- Hustota pravděpodobnosti:

$$t(n) = \frac{N(0,1)}{\chi^2(n)/n}$$

- Střední hodnota: $E[X] = 0$
- Rozptyl: $D[X] = \frac{n}{n-2}$

Exponenciální rozdělení = $Exp(\lambda)$



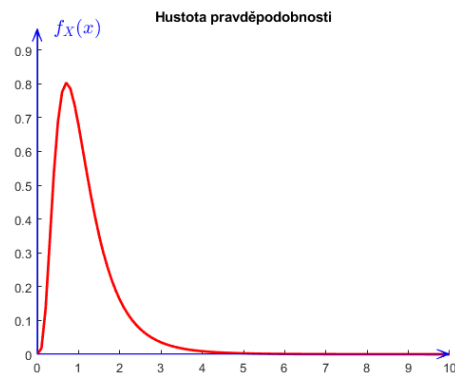
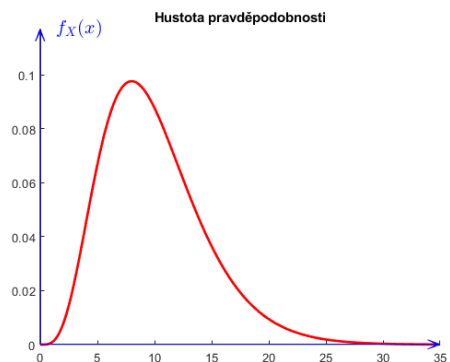
- Vyjadřuje dobu mezi náhodně se vyskytujícími událostmi.
- Spojitá obdoba geometrického rozdělení
- Hustota pravděpodobnosti:

$$f(x) = \lambda e^{-\lambda x} \quad \text{kde } x \geq 0, \lambda > 0$$

- Střední hodnota: $E[X] = \frac{1}{\lambda}$
- Rozptyl: $D[X] = \frac{1}{\lambda^2}$

χ^2 -rozdělení = $\chi^2(n)$

Fisherovo rozdělení = $F(n_1, n_2)$



- Obě rozdělení jsou umělá, tedy vychází z jiného rozdělení.
- Používají se pro testy hypotéz.
- χ^2 -rozdělení

$$\chi^2(n) = \sum_{i=1}^n (N_i(0,1))^2$$

- Fisherovo rozdělení

$$F(n) = \frac{\chi_1^2(m)/m}{\chi_2^2(n)/n}$$