Základy pravděpodobnosti
a matematické statistiky

3. Náhodná veličina a její charakteristiky

FAKULTA STROJNÍ ČVUT V PRAZE

prof. RNDr. Gejza Dohnal, CSc.

ak. rok 2023/2024

3. Náhodná veličina a její charakteristiky

Co máme?

1. elementární náhodné jevy Ω

2. jevové pole $\mathcal{F}$: (1) $\emptyset \in \mathcal{F}$

(2) $A \in \mathcal{F} \Rightarrow \bar{A} \in \mathcal{F}$

(3) $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

3. Pravděpodobnost (a) $\forall A \in \mathcal{F}: \quad P(A) \in \langle 0, 1 \rangle$

(b) $P(\emptyset) = 0, \quad P(\Omega) = 1$

(c) jsou-li $A_1, A_2, \dots \in \mathcal{F}$, po dvou disjunktní, potom

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

$\Rightarrow \quad (\Omega, \mathcal{F}, P)$ …. pravděpodobnostní prostor

Co nám chybí?

4. Náhodná veličina: $X : \Omega \to \mathbf{R}$

Přesněji: měřitelná funkce $X : (\Omega, \mathcal{F}) \to (\mathbf{R}, \mathcal{B})$

$$\forall x \in \mathbf{R} : \quad \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$$

$$P(\{\omega \in \Omega : X(\omega) \leq x\}) = F(x), \quad x \in \mathbf{R}$$

Náhodná veličina

$$X : (\Omega, \mathcal{F}) \to (\mathbf{R}, \mathcal{B})$$

$$\{\omega \in \Omega : X(\omega) \in (-\infty, x\rangle \} \in \mathcal{F}$$

$$\{X \le x\} \in \mathcal{F}$$

Distribuční funkce náhodné veličiny: $F(x) = P(X \le x)\,, \quad x \in \mathbf{R}$

Distribuční funkce odpovídá na otázku: S jakou pravděpodobností náhodná veličina nepřekročí hodnotu x?

Rozlišujeme dva případy:

Náhodná veličina je funkce z Ω do

diskrétní náhodná veličina

a) spočetné podmnožiny $S \subset \mathbf{R}$

b) souvislé podmnožiny $Q \subset \mathbf{R}$

spojitá náhodná veličina

Náhodná veličina

$$X : (\Omega, \mathcal{F}) \to (\mathbf{R}, \mathcal{B})$$

Rozlišujeme dva případy:

Náhodná veličina je funkce z Ω do

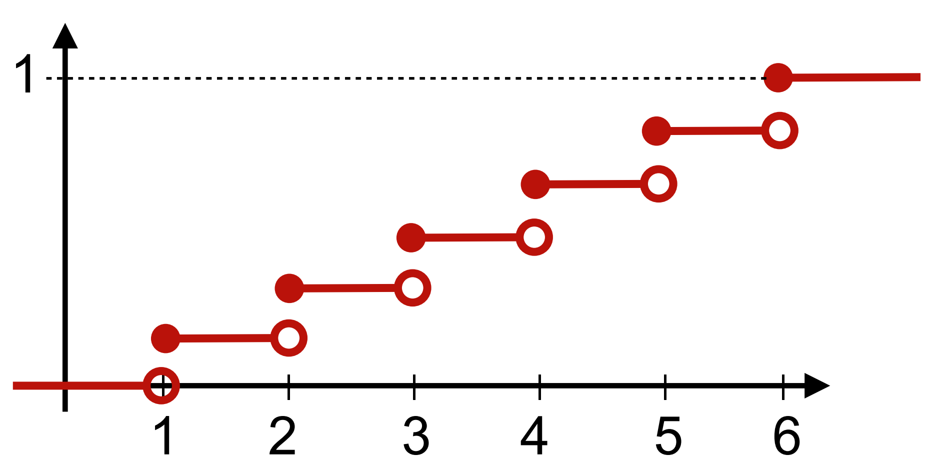
a) spočetné podmnožiny $S \subset \mathbf{R}$ — diskrétní náhodná veličina

b) souvislé podmnožiny $Q \subset \mathbf{R}$ — spojitá náhodná veličina

a) Diskrétní náhodná veličina:

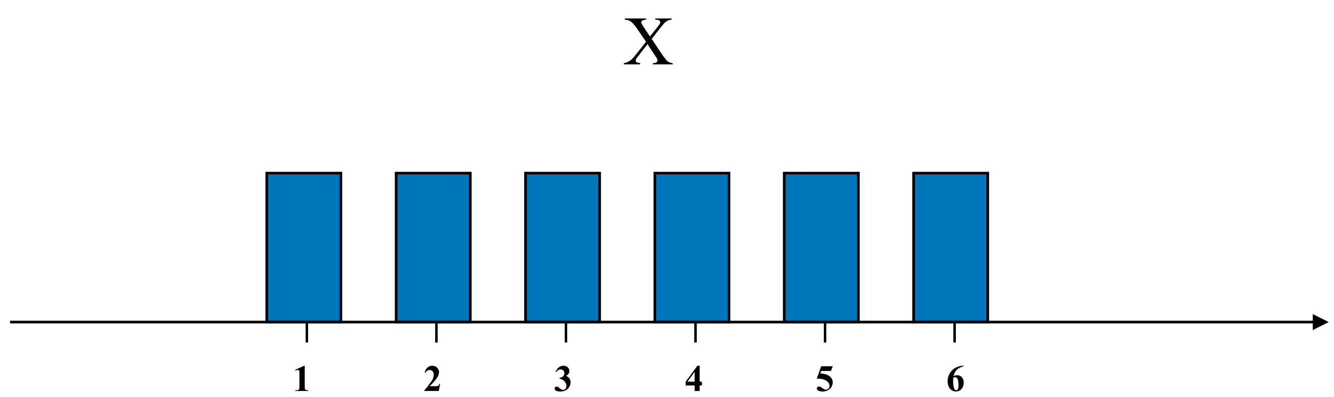
- obor hodnot = $\{x_1, x_2, \ldots, x_n, \ldots\}$
- pravděpodobnostní funkce: $P(X = x_i) = p_i, \; i=1, 2, \ldots$

  vždycky musí platit, že $p_i \in \langle 0, 1 \rangle, \quad \sum_{i=1}^{\infty} p_i = 1$
- distribuční funkce: $F(x) = P(X \leq x) = \sum_{i=:x_i \leq x} p_i, \quad x \in \mathbf{R}$



Rozdělení pravděpodobnosti náhodné veličiny

- Diskrétní náhodná veličina, pravděpodobnostní funkce



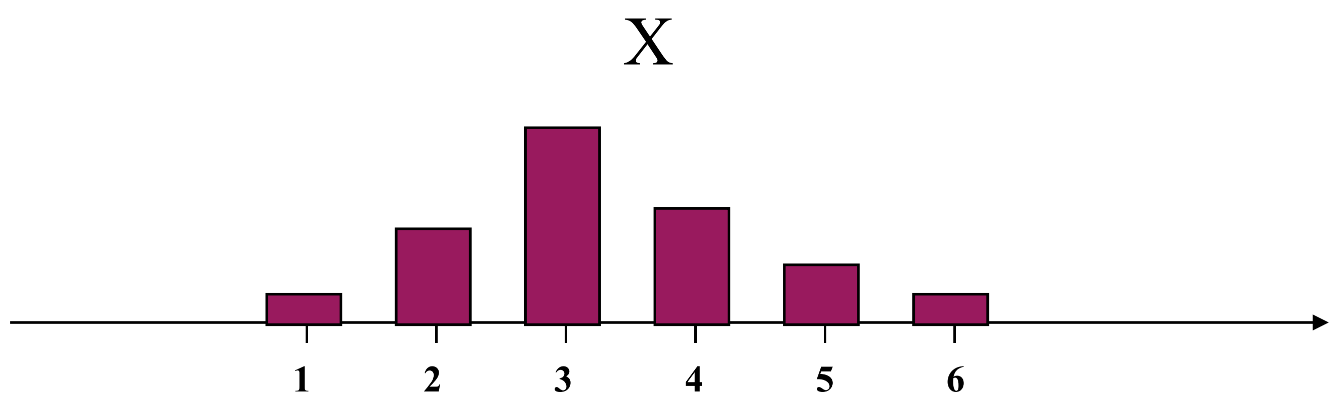
$P(X=1)= 1/6$
$P(X=2)= 1/6$
$P(X=3)= 1/6$
$P(X=4)= 1/6$
$P(X=5)= 1/6$
$P(X=6)= 1/6$

$\overline{}$
1

rozdělení pravděpodobnosti

Rozdělení pravděpodobnosti náhodné veličiny

- Diskrétní náhodná veličina, pravděpodobnostní funkce



$$
\begin{aligned}
P(X=1) &= 1/20\\
P(X=2) &= 1/6\\
P(X=3) &= 1/3\\
P(X=4) &= 1/4\\
P(X=5) &= 2/15\\
P(X=6) &= \underline{1/15}\\
& \quad 1
\end{aligned}
$$

rozdělení pravděpodobnosti

Rozdělení pravděpodobnosti náhodné veličiny

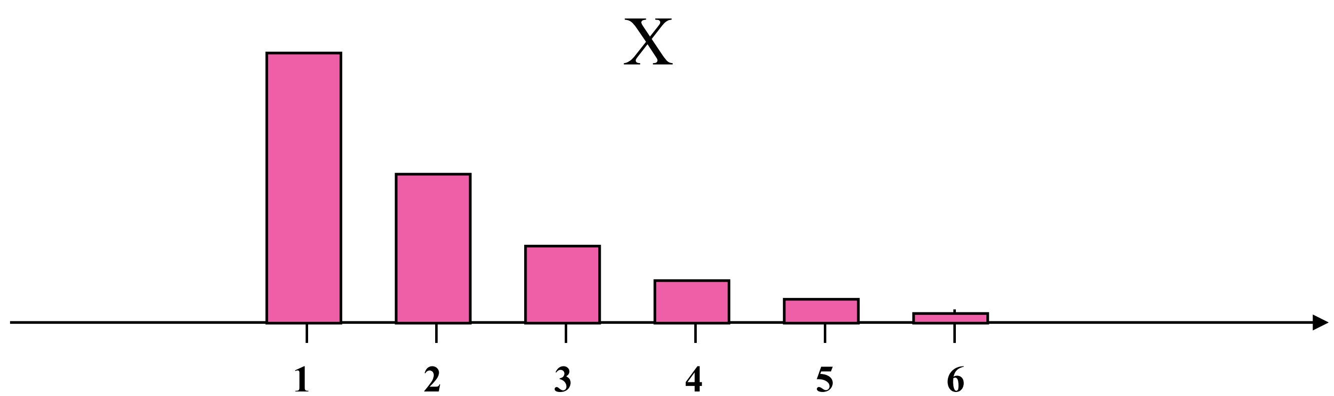
Pro distribuční funkci platí:

1) je neklesající $x \le y \;\Rightarrow \{X \le x\} \subset \{X \le y\} \Rightarrow P(X \le x) \le P(X \le y) \Rightarrow F(x) \le F(y).$

2) $\lim\limits_{x \to -\infty} F(x) = \lim\limits_{x \to -\infty} P(X \le x) = 0$; $\lim\limits_{x \to +\infty} F(x) = \lim\limits_{x \to +\infty} P(X \le x) = 1$

3) $F(x)$ je zprava spojitá $\lim\limits_{x \to x_0+} F(x) = F(x_0)$

- **Diskrétní náhodná veličina, pravděpodobnostní funkce**

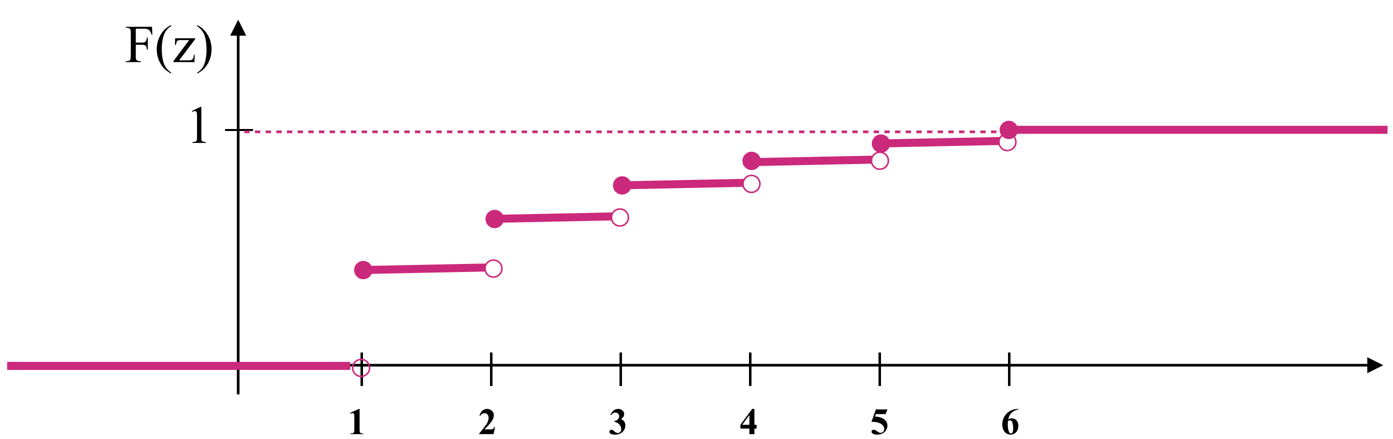


P(X=1)= 32/55
P(X=2)= 16/55
P(X=3)= 8/55
P(X=4)= 4/55
P(X=5)= 2/55
P(X=6)= 1/55

1

rozdělení pravděpodobnosti

Distribuční funkce: $F_X(z) = P(X \le z), \quad -\infty < z < +\infty$



Rozdělení pravděpodobnosti náhodné veličiny

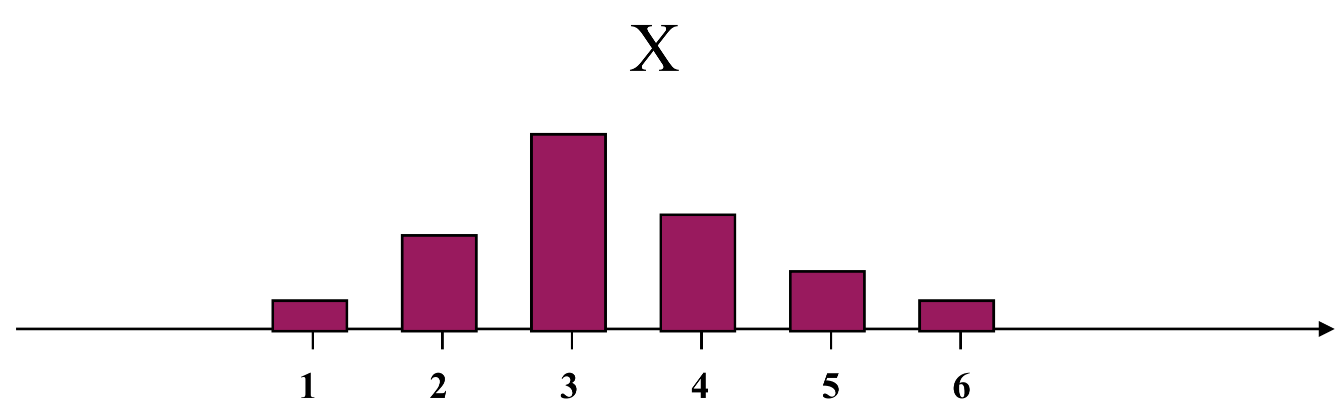
Pro distribuční funkci platí:

1) je neklesající $x \leq y \;\Rightarrow \{X \leq x\} \subset \{X \leq y\} \Rightarrow P(X \leq x) \leq P(X \leq y) \Rightarrow F(x) \leq F(y).$

2) $\lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} P(X \leq x) = 0 \;;\; \lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} P(X \leq x) = 1$

3) $F(x)$ je zprava spojitá $\lim_{x \to x_0+} F(x) = F(x_0)$

- **Diskrétní náhodná veličina, pravděpodobnostní funkce**

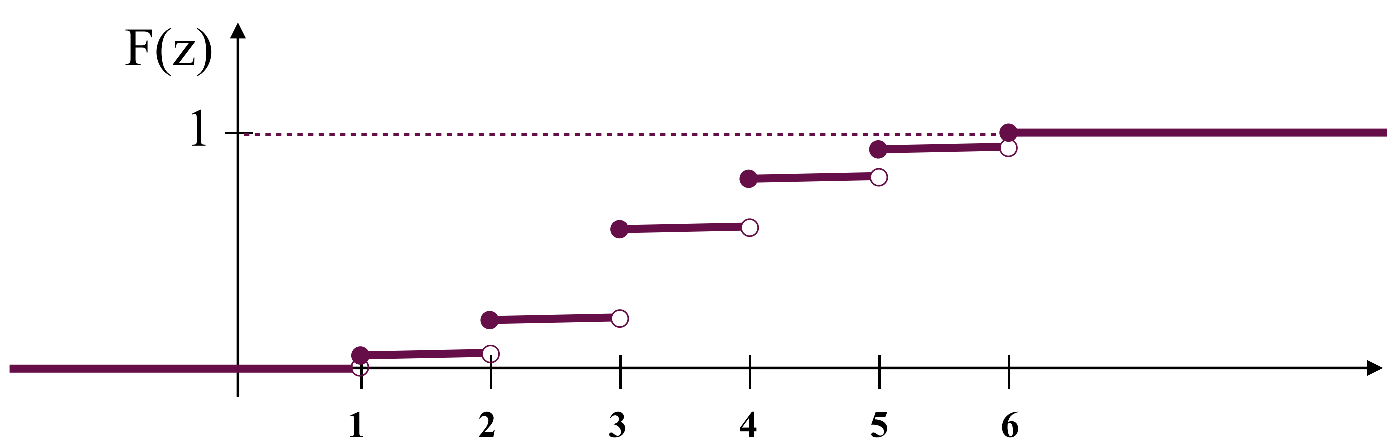


P(X=1)= 1/20
P(X=2)= 1/6
P(X=3)= 1/3
P(X=4)= 1/4
P(X=5)= 2/15
P(X=6)= 1/15

1

rozdělení pravděpodobnosti

Distribuční funkce: $F_X(z) = P(X \leq z), \quad -\infty < z < +\infty$



Rozdělení pravděpodobnosti náhodné veličiny

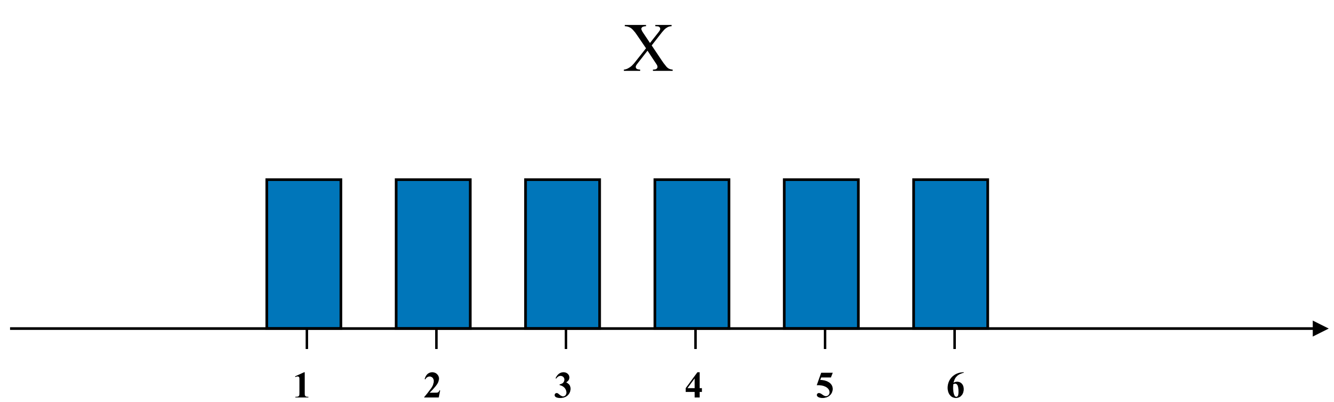
Pro distribuční funkci platí:

1) je neklesající $x \leq y \; \Rightarrow \{X \leq x\} \subset \{X \leq y\} \Rightarrow P(X \leq x) \leq P(X \leq y) \Rightarrow F(x) \leq F(y).$

2) $\lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} P(X \leq x) = 0 \; ; \; \lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} P(X \leq x) = 1$

3) $F(x)$ je zprava spojitá $\lim_{x \to x_0+} F(x) = F(x_0)$

- **Diskrétní náhodná veličina, pravděpodobnostní funkce**

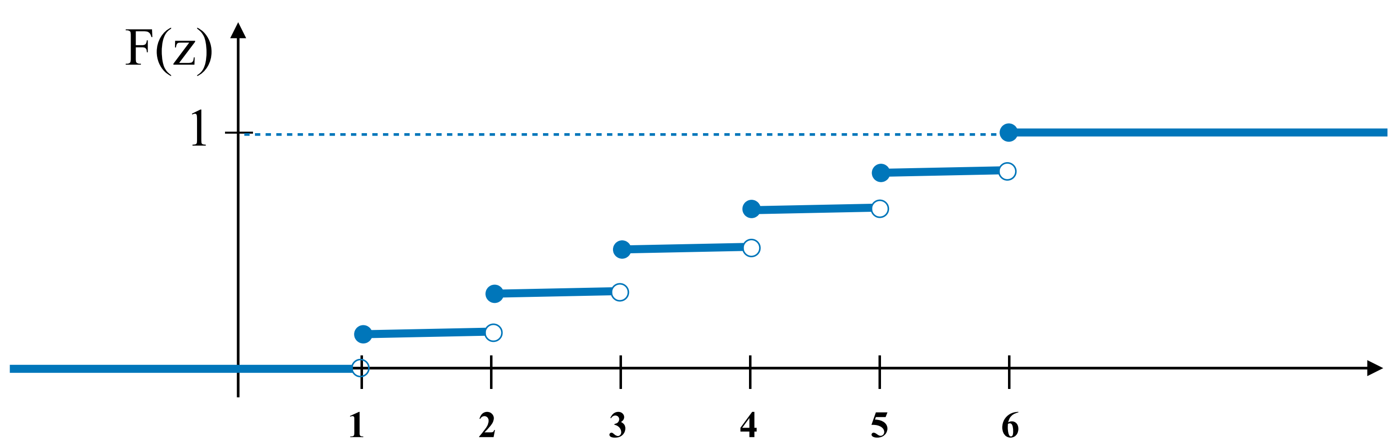


P(X=1)= 1/6
P(X=2)= 1/6
P(X=3)= 1/6
P(X=4)= 1/6
P(X=5)= 1/6
P(X=6)= 1/6

1

rozdělení pravděpodobnosti

Distribuční funkce: $F_X(z) = P(X \leq z), \quad -\infty < z < +\infty$



Rozdělení pravděpodobnosti náhodné veličiny

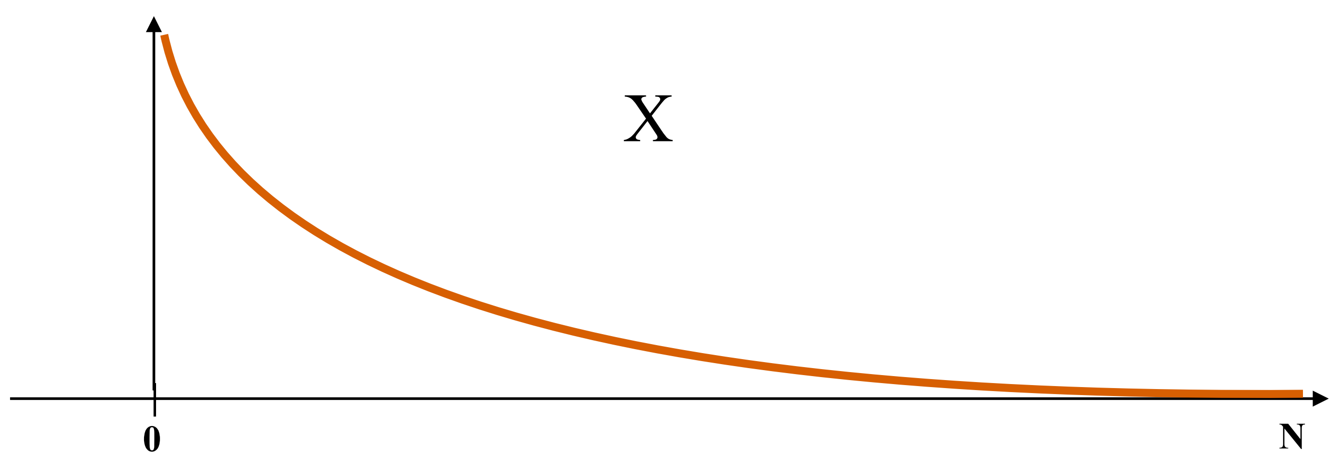
Pro distribuční funkci platí:

1) je neklesající $x \le y \;\Rightarrow \{X \le x\} \subset \{X \le y\} \Rightarrow P(X \le x) \le P(X \le y) \Rightarrow F(x) \le F(y).$

2) $\lim\limits_{x \to -\infty} F(x) = \lim\limits_{x \to -\infty} P(X \le x) = 0 \;;\; \lim\limits_{x \to +\infty} F(x) = \lim\limits_{x \to +\infty} P(X \le x) = 1$

3) $F(x)$ je zprava spojitá $\lim\limits_{x \to x_0+} F(x) = F(x_0)$

- **Diskrétní náhodná veličina, pravděpodobnostní funkce**

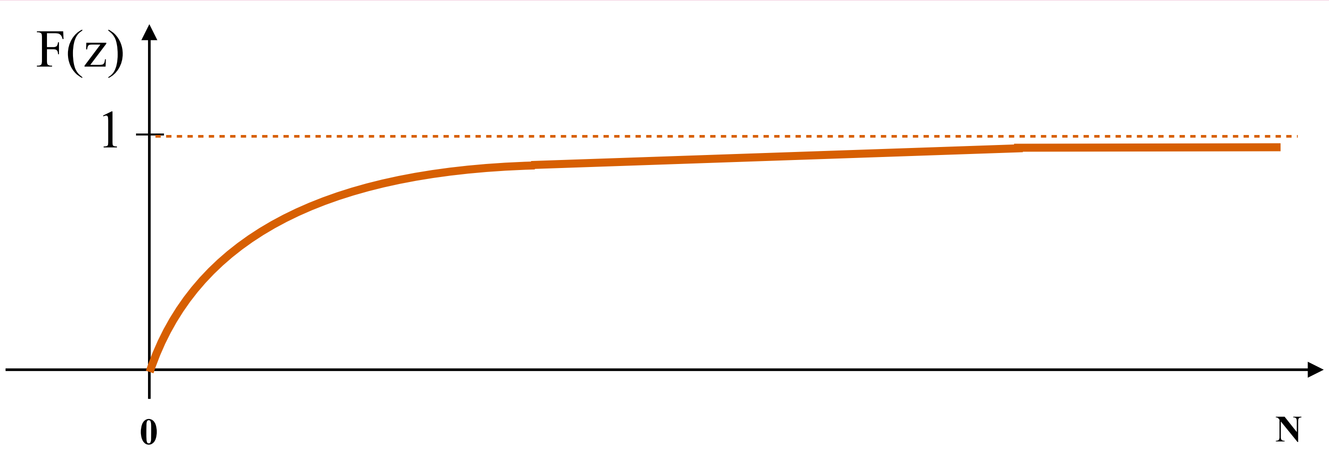


$$P(X = k) = p(1-p)^{k-1}$$

$$\sum_{k=1}^{\infty} P(X = k) = 1$$

rozdělení pravděpodobnosti

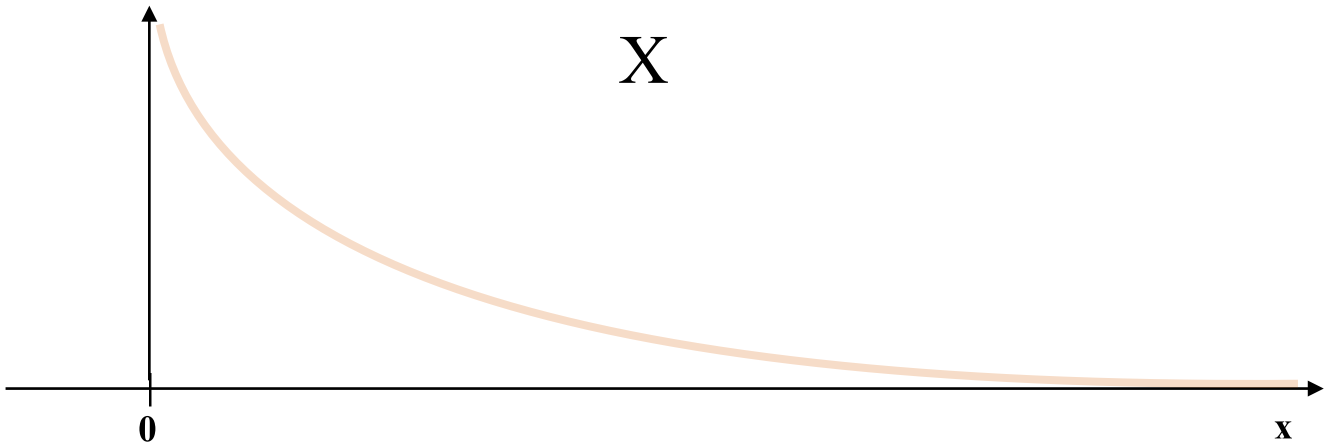
Distribuční funkce: $F_X(z) = P(X \le z), \quad -\infty < z < +\infty$



Rozdělení pravděpodobnosti náhodné veličiny

- **Spojitá náhodná veličina, funkce hustoty**

$$P(X = x) = 0$$

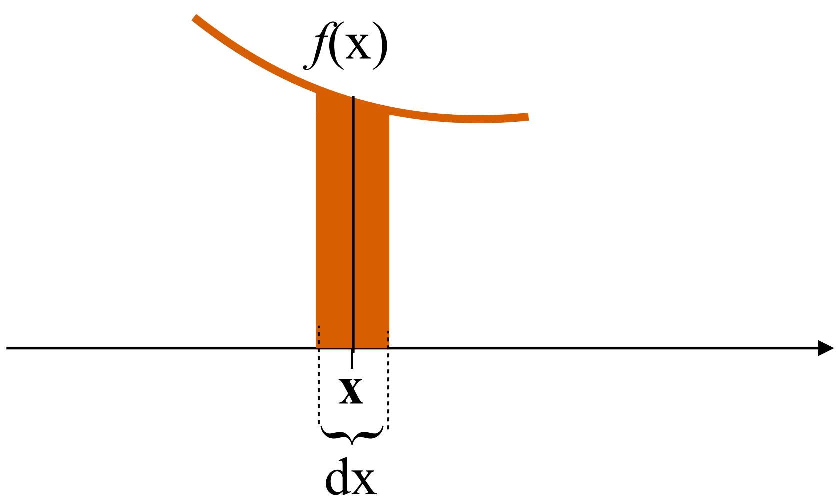


Rozdělení pravděpodobnosti náhodné veličiny

- **Spojitá náhodná veličina, funkce hustoty** $P(X = x) = 0$

hustota pravděpodobnosti: spojitá nezáporná funkce $f(x)$

$$P(X \in \langle\, x - \mathrm{d}x\,;\, x + \mathrm{d}x\,\rangle) = f\,(x)\mathrm{d}x$$



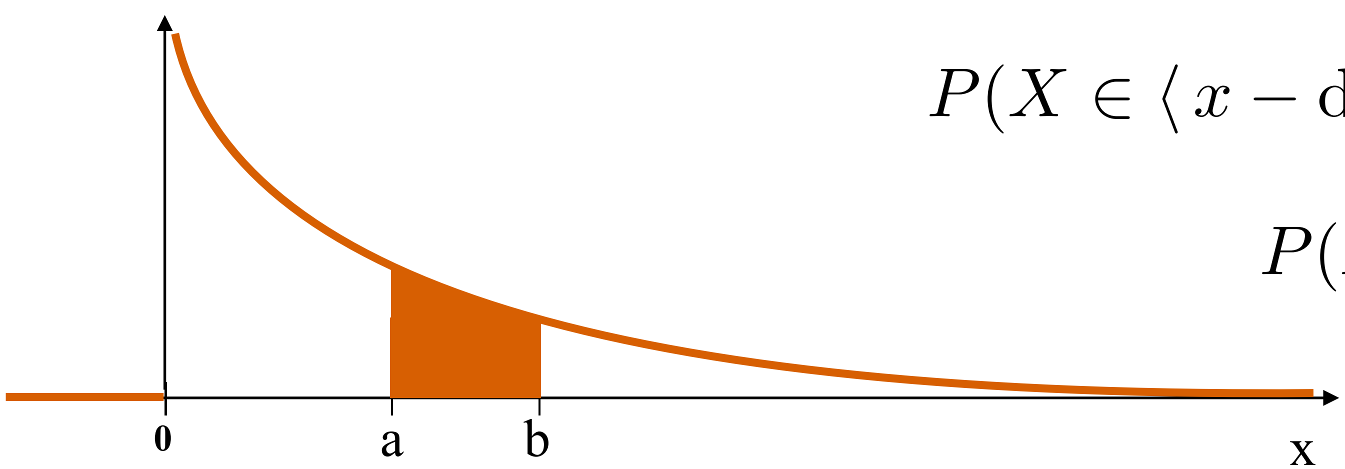
rozdělení pravděpodobnosti

Rozdělení pravděpodobnosti náhodné veličiny

- **Spojitá náhodná veličina, funkce hustoty** $P(X = x) = 0$

rozdělení pravděpodobnosti

hustota pravděpodobnosti: spojitá nezáporná funkce $f(x)$



$$P(X \in \langle x - \mathrm{d}x\,;\, x + \mathrm{d}x\,\rangle) = f(x)\mathrm{d}x$$

$$P(X \in \langle\, a\,;\, b\,\rangle) = \int_a^b f(x)\mathrm{d}x$$

$$\int_{x=-\infty}^{\infty} f(x)\mathrm{d}x = 1$$

Distribuční funkce: $F_X(x) = P(X \leq x)$, $-\infty < x < +\infty$ $\qquad F_X(x) = \int_{x=-\infty}^{x} f(z)\mathrm{d}z$

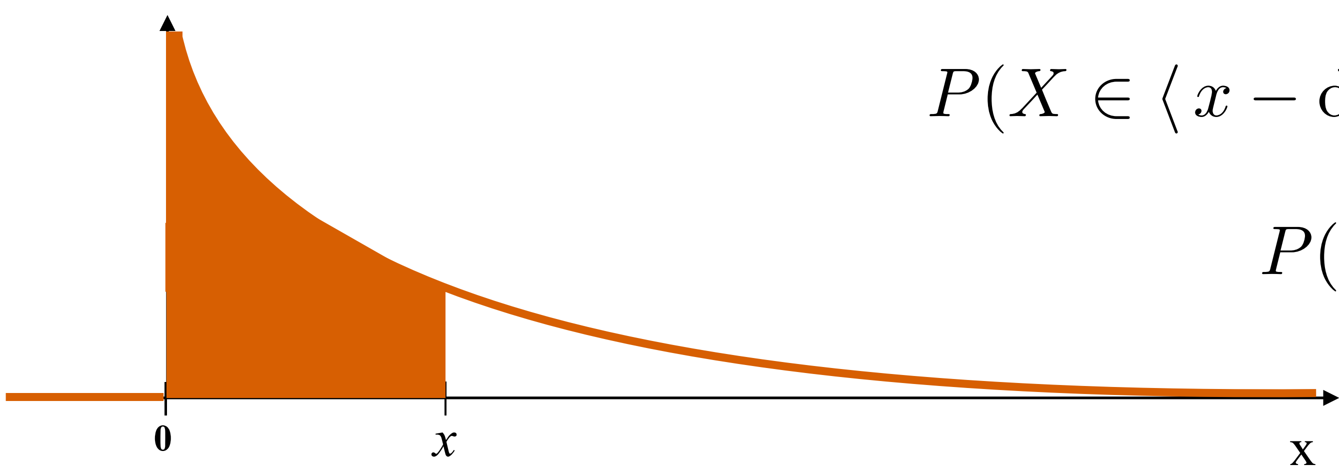
Rozdělení pravděpodobnosti náhodné veličiny

- **Spojitá náhodná veličina, funkce hustoty**

$P(X = x) = 0$

rozdělení pravděpodobnosti

hustota pravděpodobnosti: spojitá nezáporná funkce $f(x)$



$$P(X \in \langle x - \mathrm{d}x \,;\, x + \mathrm{d}x \,\rangle) = f(x)\mathrm{d}x$$

$$P(X \in \langle a \,;\, b \rangle) = \int_a^b f(x)\mathrm{d}x$$

$$\int_{x=-\infty}^{\infty} f(x)\mathrm{d}x = 1$$

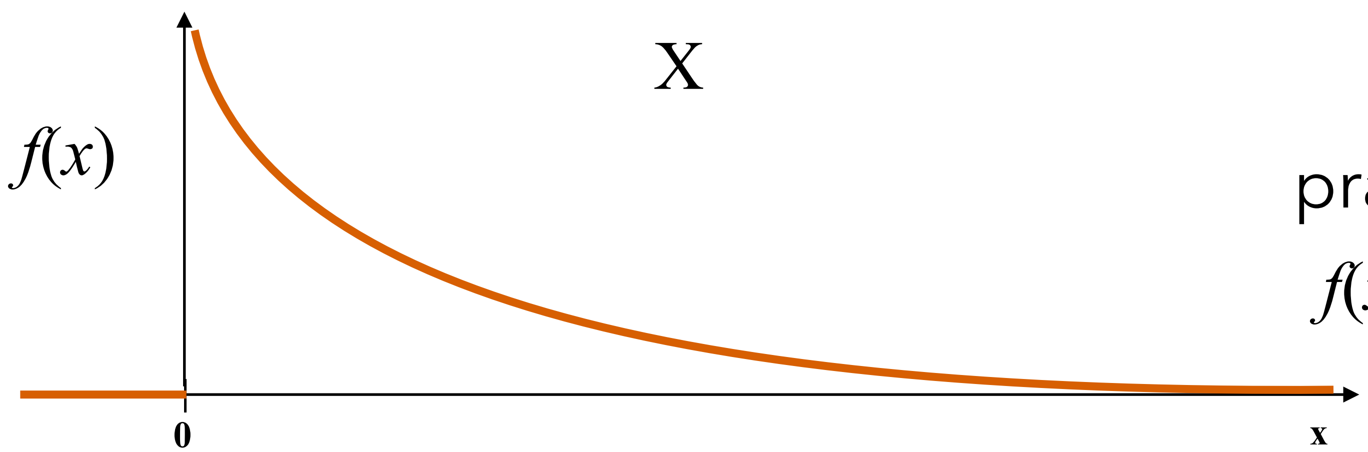
Distribuční funkce: $F_X(x) = P(X \leq x)$, $-\infty < x < +\infty$

$$F_X(x) = \int_{x=-\infty}^{x} f(z)\mathrm{d}z$$

$$f(x) = \frac{\mathrm{d}F_X}{\mathrm{d}x}$$

Rozdělení pravděpodobnosti náhodné veličiny

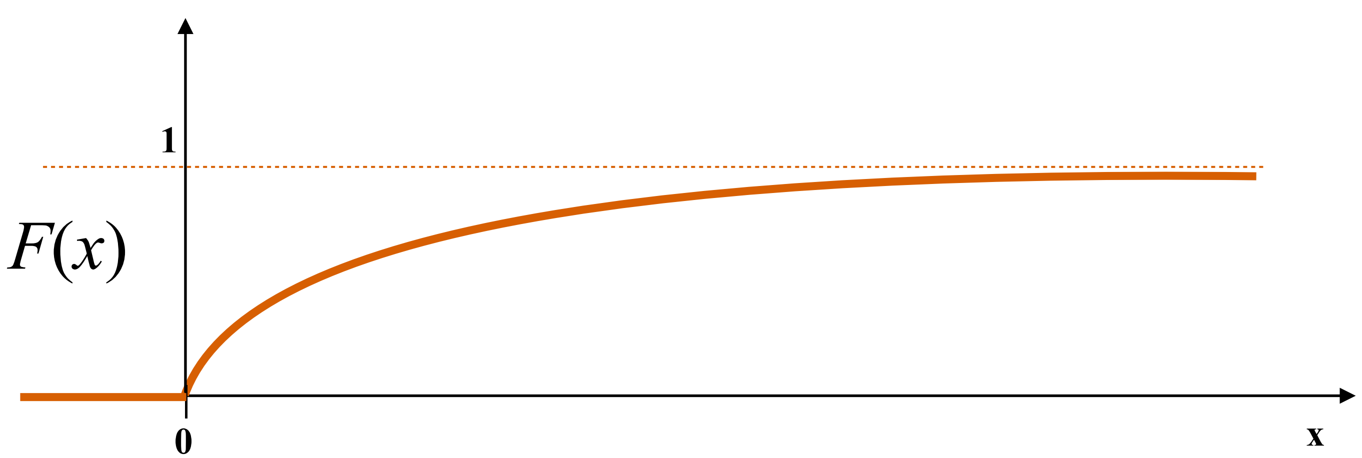
- **Spojitá náhodná veličina, funkce hustoty**



hustota pravděpodobnosti

$f(x)$, $-\infty < x < +\infty$

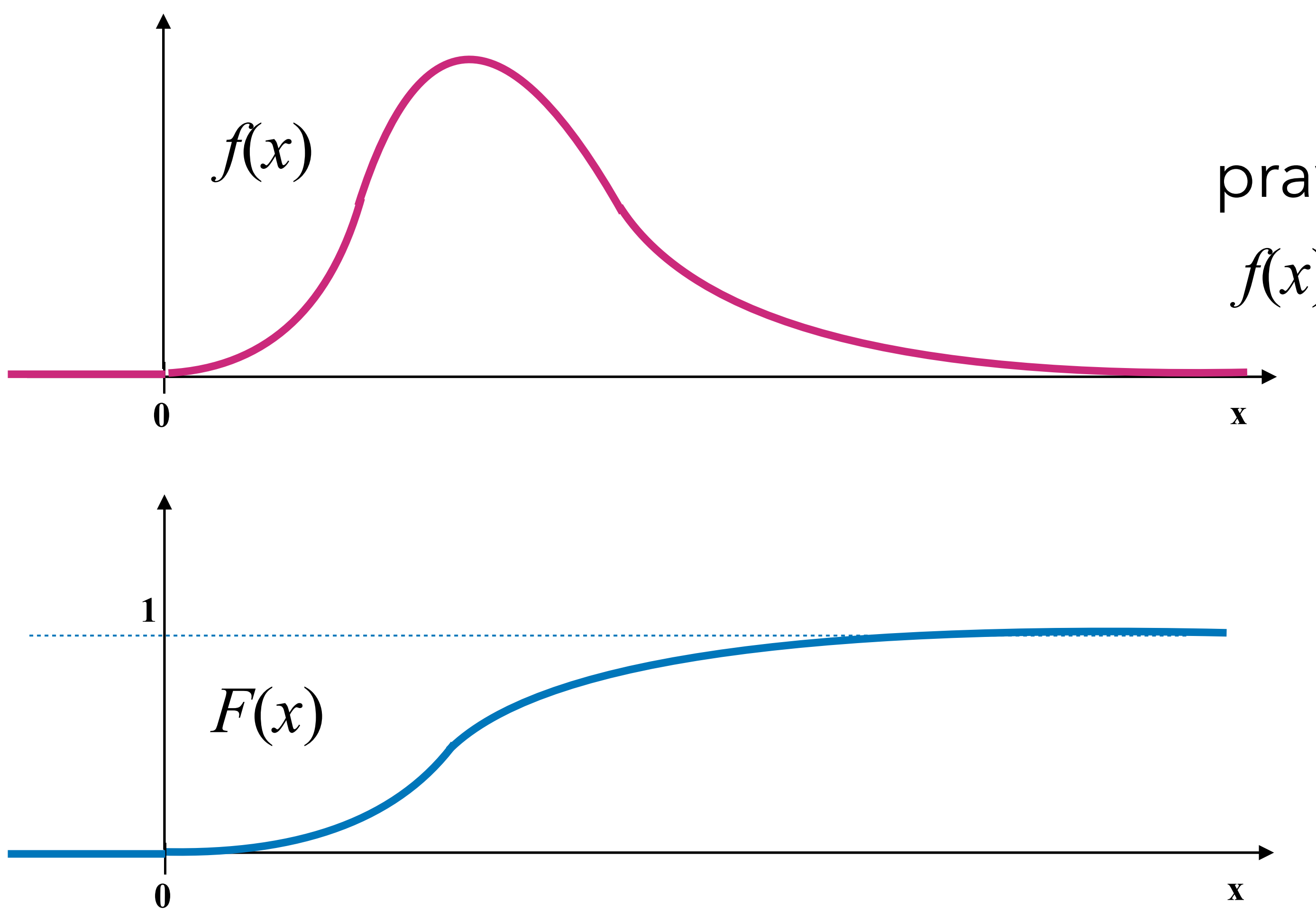
rozdělení pravděpodobnosti



$$F_X(x) = \int_{x=-\infty}^{x} f(z)\mathrm{d}z$$

Rozdělení pravděpodobnosti náhodné veličiny

- **Spojitá náhodná veličina, funkce hustoty**



hustota pravděpodobnosti

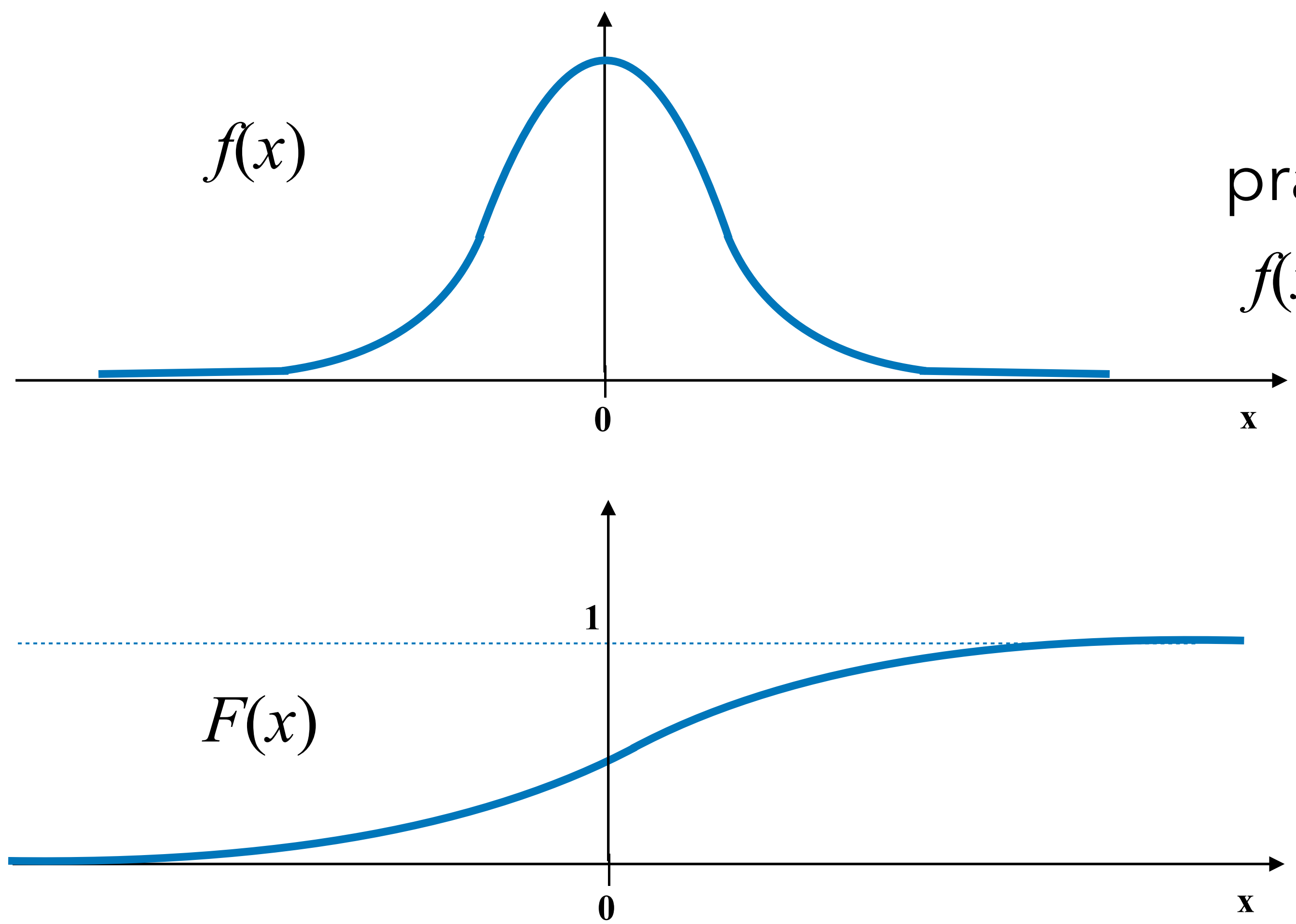
$f(x)$, $-\infty < x < +\infty$

$$F_X(x) = \int_{x=-\infty}^{x} f(z)\mathrm{d}z$$

rozdělení pravděpodobnosti

Rozdělení pravděpodobnosti náhodné veličiny

- **Spojitá náhodná veličina, funkce hustoty**



hustota pravděpodobnosti

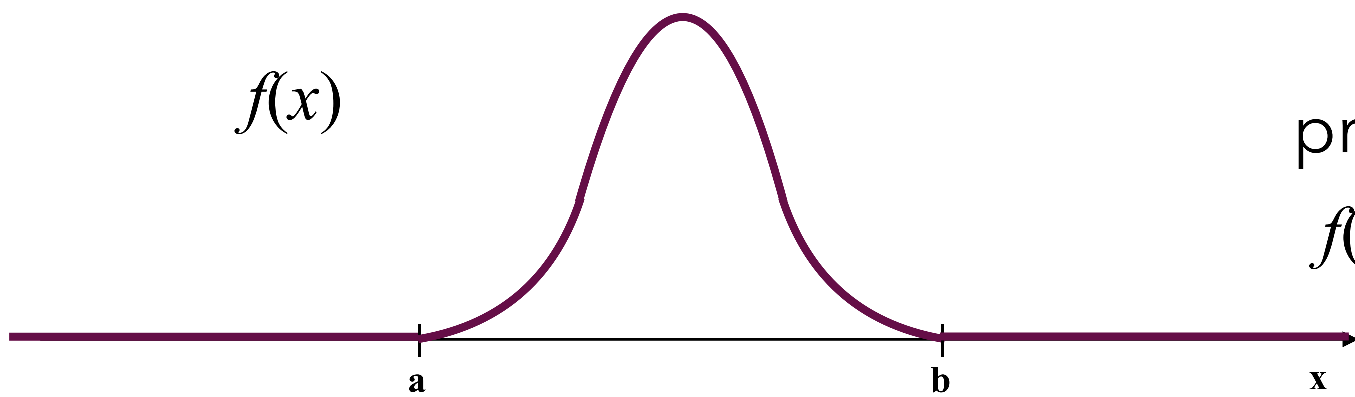
$f(x), \; -\infty < x < +\infty$

rozdělení pravděpodobnosti

$$F_X(x) = \int_{x=-\infty}^{x} f(z)\mathrm{d}z$$

Rozdělení pravděpodobnosti náhodné veličiny

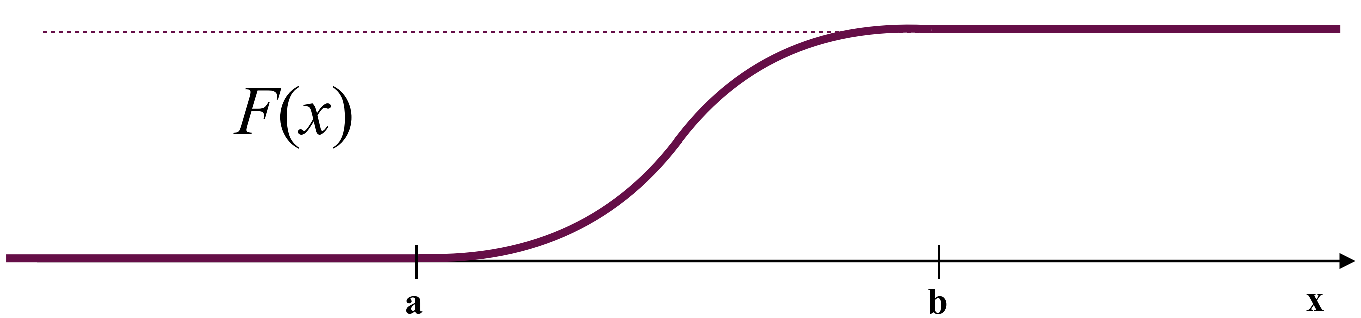
- **Spojitá náhodná veličina, funkce hustoty**



hustota pravděpodobnosti

$f(x)$, $-\infty < x < +\infty$

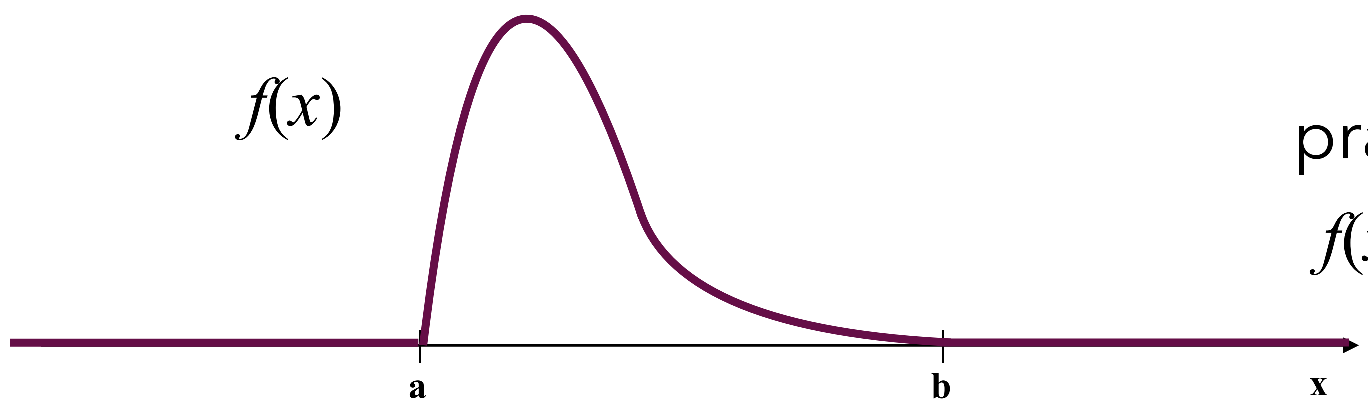
rozdělení pravděpodobnosti



$$F_X(x) = \int_{x=-\infty}^{x} f(z)\,\mathrm{d}z$$

Rozdělení pravděpodobnosti náhodné veličiny

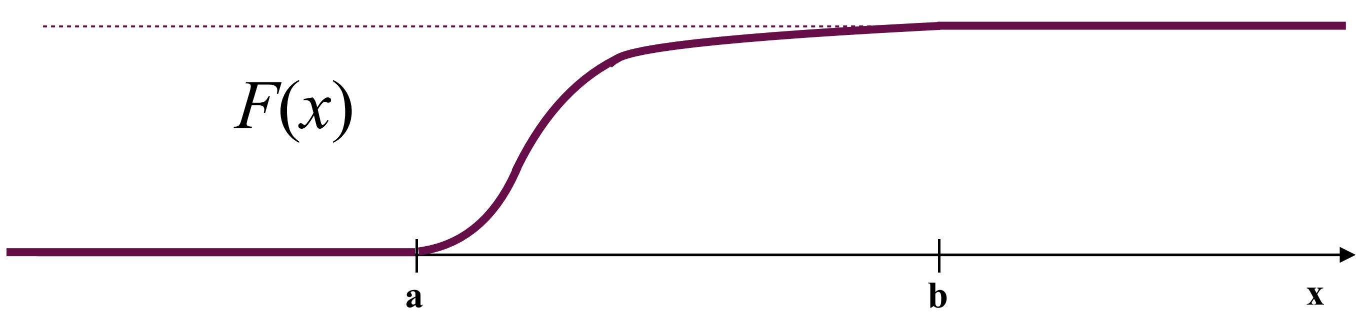
- **Spojitá náhodná veličina, funkce hustoty**



hustota pravděpodobnosti

$f(x)$, $-\infty < x < +\infty$

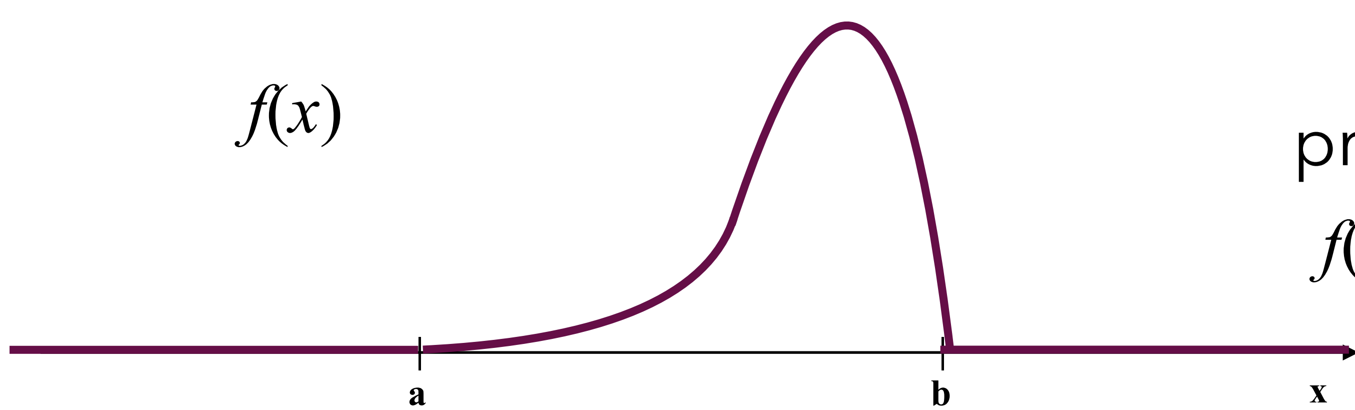
rozdělení pravděpodobnosti



$$F_X(x) = \int_{x=-\infty}^{x} f(z)\mathrm{d}z$$

Rozdělení pravděpodobnosti náhodné veličiny

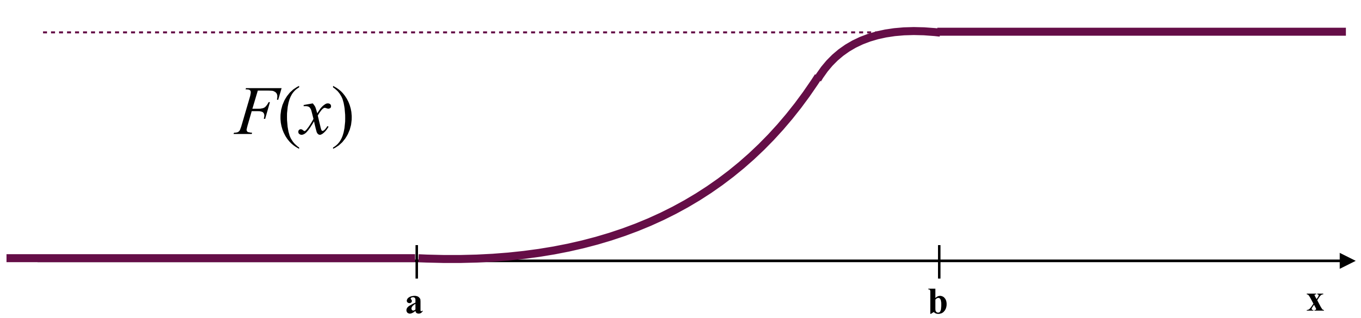
- **Spojitá náhodná veličina, funkce hustoty**



hustota pravděpodobnosti

$f(x), \; -\infty < x < +\infty$

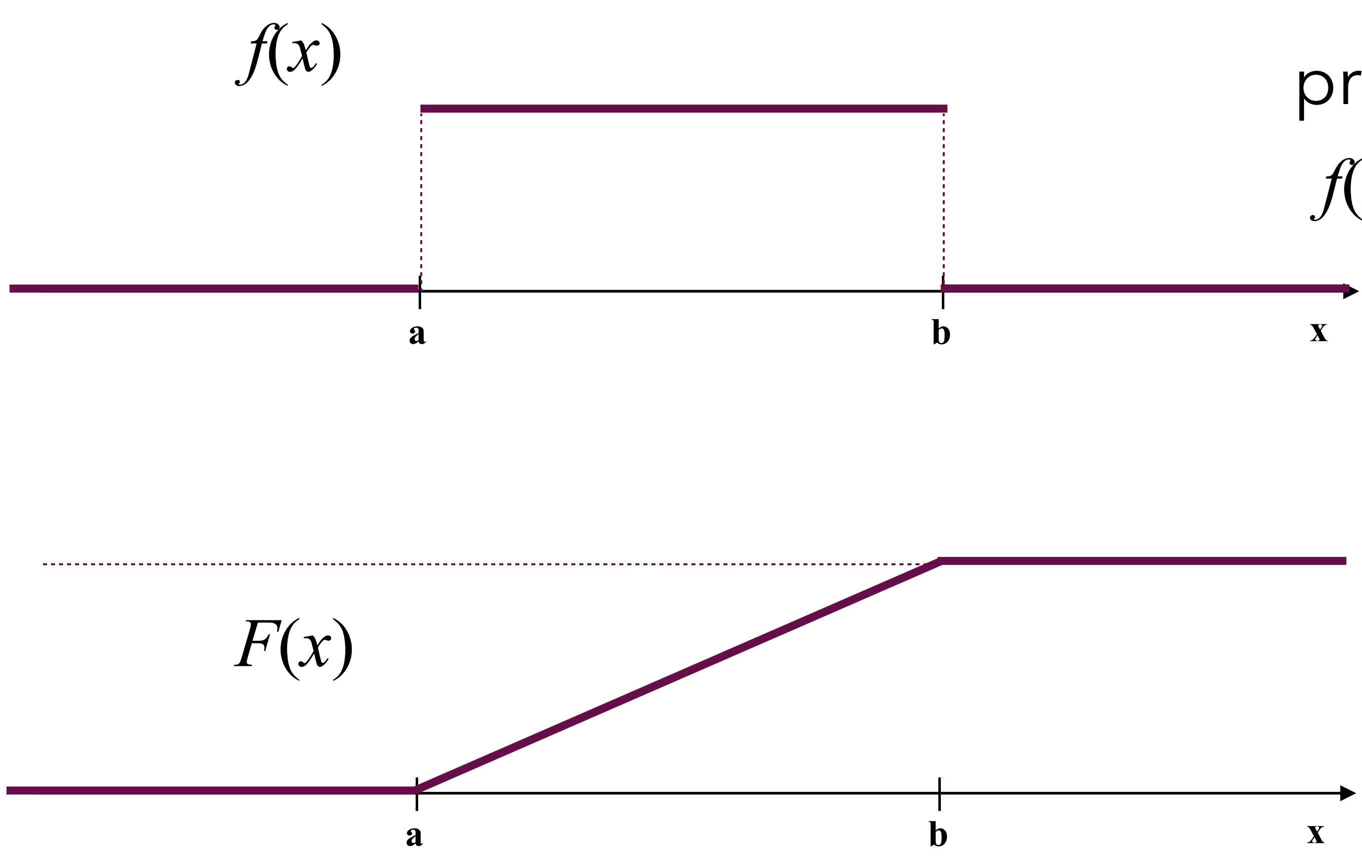
rozdělení pravděpodobnosti



$$F_X(x) = \int_{x=-\infty}^{x} f(z)\,\mathrm{d}z$$

Rozdělení pravděpodobnosti náhodné veličiny

- **Spojitá náhodná veličina, funkce hustoty**



hustota pravděpodobnosti

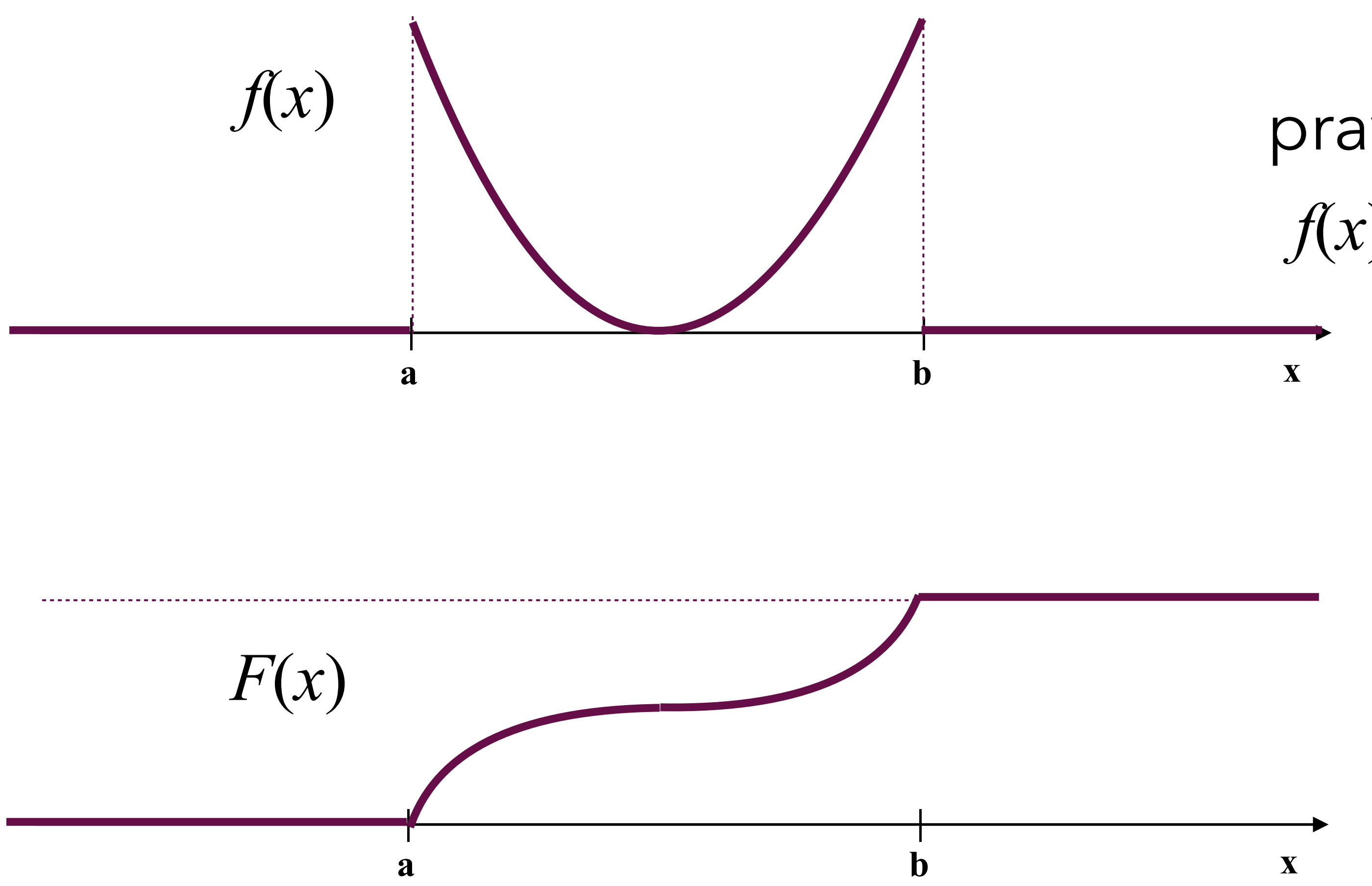
$f(x), \quad -\infty < x < +\infty$

rozdělení pravděpodobnosti

$$F_X(x) = \int_{x=-\infty}^{x} f(z)\,\mathrm{d}z$$

Rozdělení pravděpodobnosti náhodné veličiny

- **Spojitá náhodná veličina, funkce hustoty**



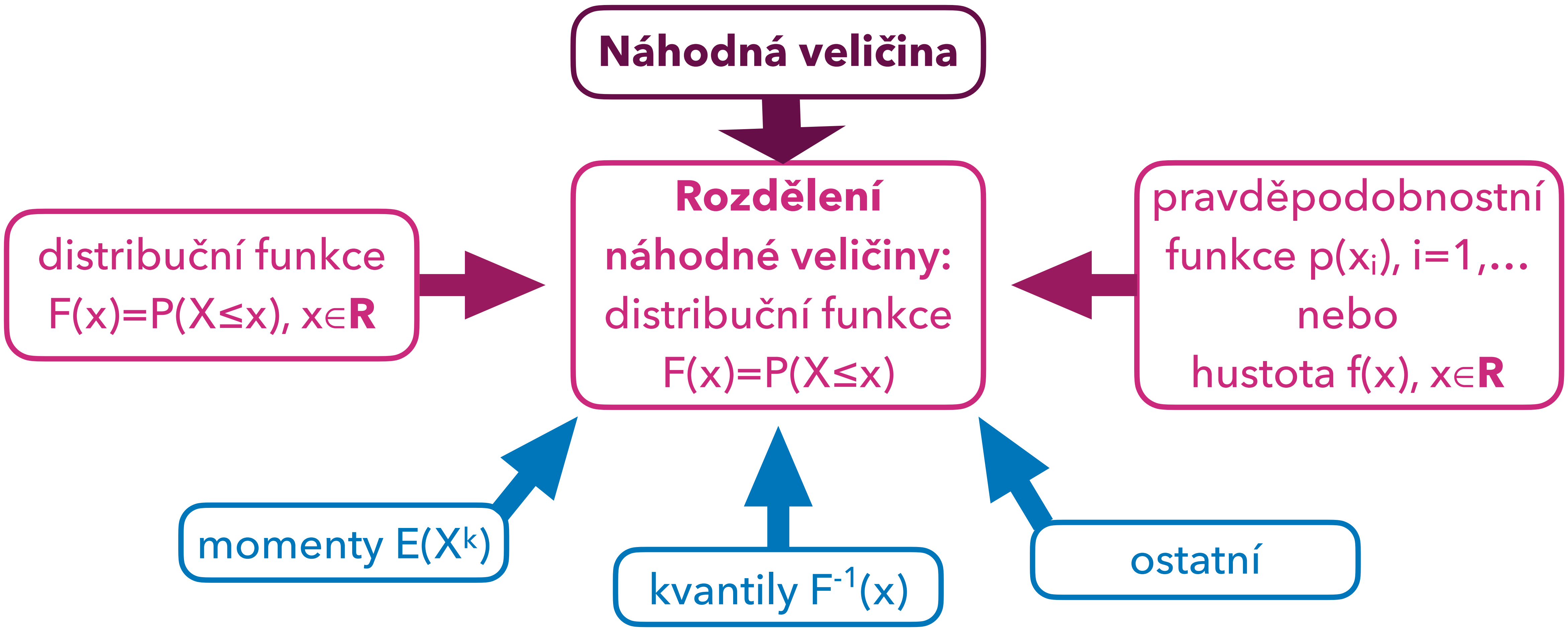
hustota pravděpodobnosti

$f(x)$, $-\infty < x < +\infty$

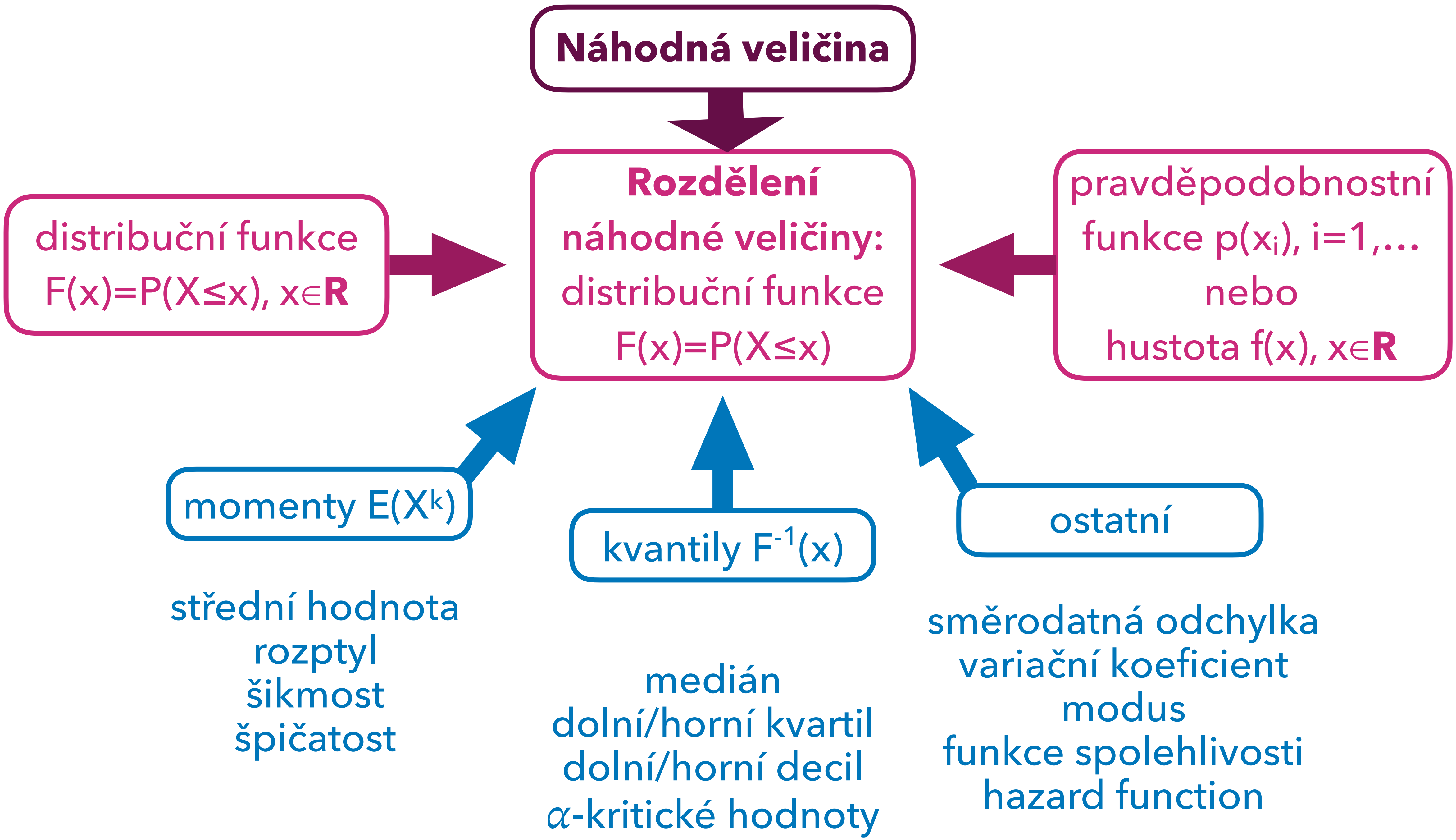
$$F_X(x) = \int_{x=-\infty}^{x} f(z)\mathrm{d}z$$

rozdělení pravděpodobnosti

Charakteristiky náhodné veličiny



Charakteristiky náhodné veličiny



Charakteristiky náhodné veličiny - Momenty

Střední hodnota: $E(X) = \int_{-\infty}^{+\infty} x \mathrm{d}F(x)$

a) v případě diskrétní náhodné veličiny $E(X) = \sum_{i=1}^{\infty} x_i p_i$

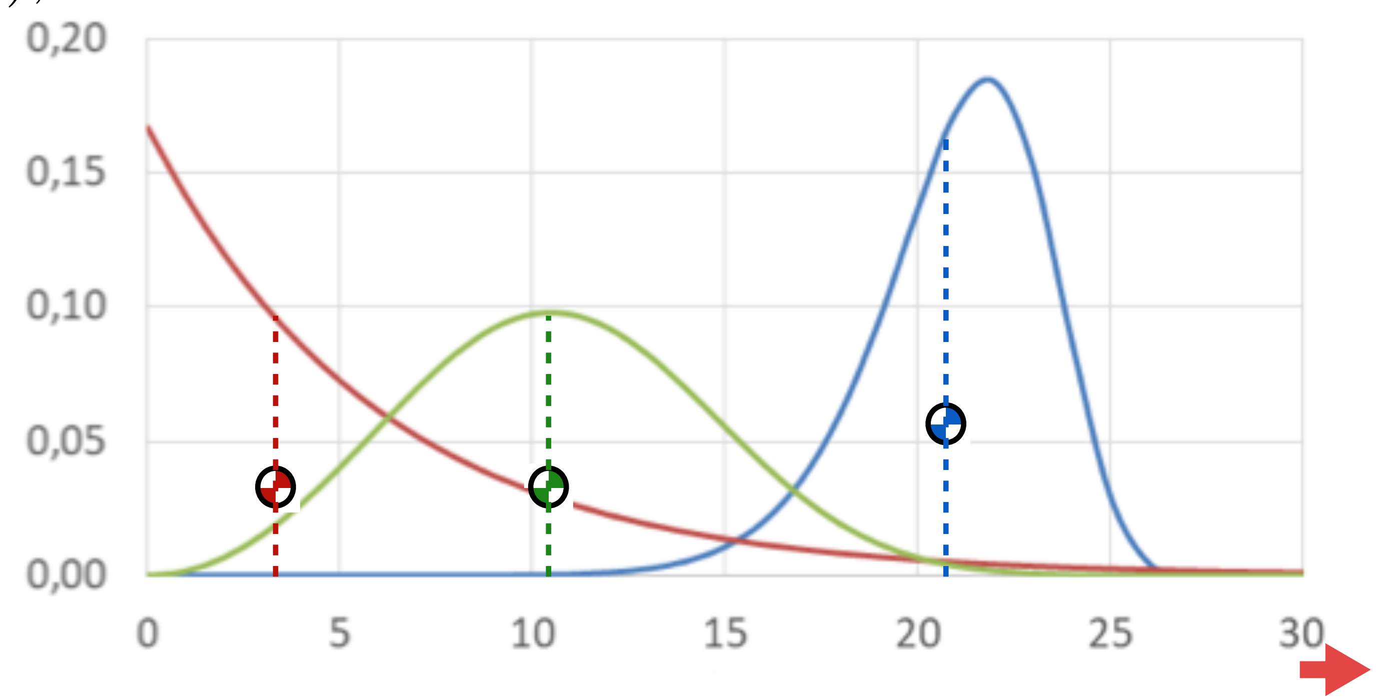
b) v případě spojité náhodné veličiny $E(X) = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x$

Střední hodnota je lineární funkcionál:

je aditivní: $E(X+Y) = E(X) + E(Y)$

a je homogenní: $E(kX) = kE(X), \quad k \in \mathbf{R}$

Střední hodnota tvoří jakési
“těžiště rozdělení“ náhodné veličiny:



Charakteristiky náhodné veličiny - Momenty

Obecné momenty: $\mu_k(X) = E(X^k) = \int_{-\infty}^{+\infty} x^k \mathrm{d}F(x)$

$\mu_k(X) = \sum_{i=1}^{\infty} x_i p_i$ v diskrétním případě (pro d.n.v.)

$\mu_k(X) = \int_{-\infty}^{+\infty} x^k f(x)\mathrm{d}x$ ve spojitém případě (pro s.n.v.)

Momentová vytvořující funkce: $M_X(t) = E(e^{tX}), \quad t \in \mathbf{R}$

$M_X(t) = \sum_{i=1}^{\infty} e^{tx_i} p_i$ pro d.n.v., $M_X(t) = \int_{-\infty}^{+\infty} e^{tx} f(x)\mathrm{d}x$ pro s.n.v.

$$M_X(t) = 1 + tE(X) + \frac{t^2}{2!}\mu_2 + \frac{t^3}{3!}\mu_3 + \cdots + \frac{t^n}{n!}\mu_n + \ldots$$

Platí například: 1) $\mu_1(X) = E(X) = \mu$ 2) $\mu_k = \frac{\mathrm{d}^k M_X}{\mathrm{d}t^k}(0)$

3) $S_n = \sum_{i=1}^{n} a_i X_i \;\Rightarrow\; M_{S_n}(t) = \prod_{i=1}^{n} M_{x_i}(a_i t)$

Charakteristiky náhodné veličiny - Momenty

Centrální momenty: $X_C = X - E(X) = X - \mu$ ….. centrovaná náhodná veličina

$$\nu_k(X) = E(X_C^k) = E(X-\mu)^k = \int_{-\infty}^{+\infty} (x-\mu)^k \mathrm{d}F(x)$$

Je zřejmě:

1) $\nu_1(X) = E(X_C) = E(X-\mu) = E(X) - \mu = E(X) - E(X) = 0$

2) $\nu_2(X) = E(X-\mu)^2 = E(X^2 - 2\mu X + \mu^2) \ = E(X^2) - 2\mu E(X) + \mu^2$

$$= E(X^2) - 2\Big(E(X)\Big)^2 + \Big(E(X)\Big)^2 = E(X^2) - \Big(E(X)\Big)^2$$

$$= \mu_2(X) - \Big(\mu_1(X)\Big)^2$$

3) $\nu_k(X) = E(X-\mu)^k = E\left(\sum_{i=0}^{k} \binom{k}{i} X^{k-i}\mu^i\right) = \sum_{i=0}^{k} \binom{k}{i} \mu_{k-i}\mu_1^i$

Charakteristiky náhodné veličiny - Momenty

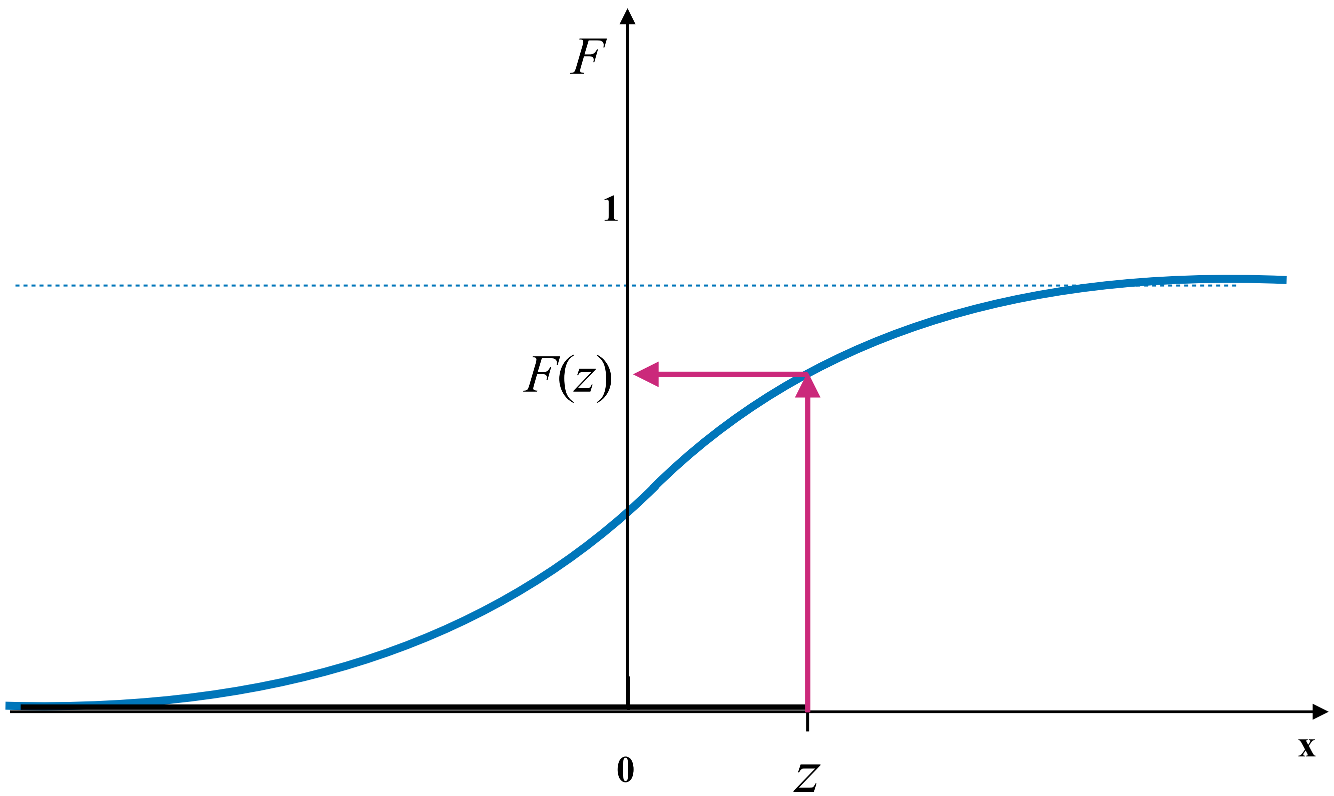
Centrální momenty: $X_C = X - E(X) = X - \mu$ ….. centrovaná náhodná veličina

$$\nu_k(X) = E(X_C^k) = E(X-\mu)^k = \int_{-\infty}^{+\infty} (x-\mu)^k \mathrm{d}F(x)$$

- druhý centrální moment = rozptyl : (míra variability) $$Var(X) = E(X - EX)^2 = \mu_2 - \mu_1^2$$
- odvozenou charakteristikou je směrodatná odchylka: $$\sigma(X) = \sqrt{Var(X)} = \sqrt{E(X-\mu)^2} = \sqrt{\nu_2}$$
- další odvozenou charakteristikou je variační koeficient: $$V = \frac{\sigma(X)}{E(X)} = \frac{\sqrt{(\nu_2)}}{\mu_1}$$
- třetí centrální moment je mírou symetrie; koeficient šikmosti: $$Skew(X) = \frac{E(X-EX)^3}{(Var(X))^{3/2}} = \frac{\nu_3}{\mu_2^{3/2}}$$
- čtvrtý centrální moment je mírou koncentrace; koeficient špičatosti: $$Kurt(X) = \frac{E(X-EX)^4}{(Var(X))^2} = \frac{\nu_4}{\mu_2^2}$$

Charakteristiky náhodné veličiny - Kvantily

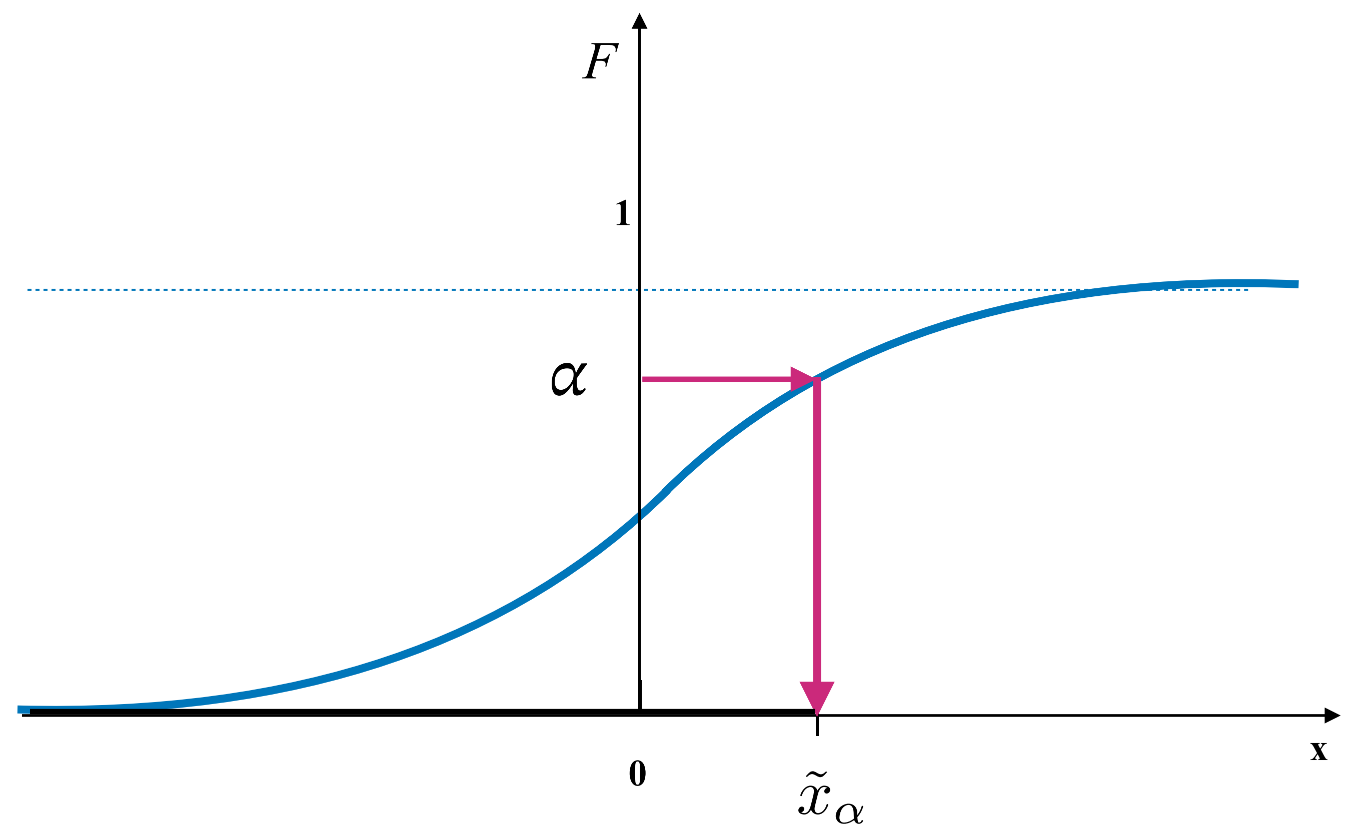
Distribuční funkce:
S jakou pravděpodobností náhodná veličina nepřekročí hodnotu z?



Příklad: Je-li X doba do poruchy součástky (lomu), potom distribuční funkce říká, s jakou pravděpodobností se součástka porouchá do doby z.

Často se však ptáme: Jakou hodnotu sledovaná náhodná veličina nepřekročí s danou pravděpodobností α?

Charakteristiky náhodné veličiny - Kvantily



Distribuční funkce:
S jakou pravděpodobností náhodná veličina nepřekročí hodnotu z?

Příklad: Je-li X doba do poruchy součástky (lomu), potom distribuční funkce říká, s jakou pravděpodobností se součástka porouchá do doby z.

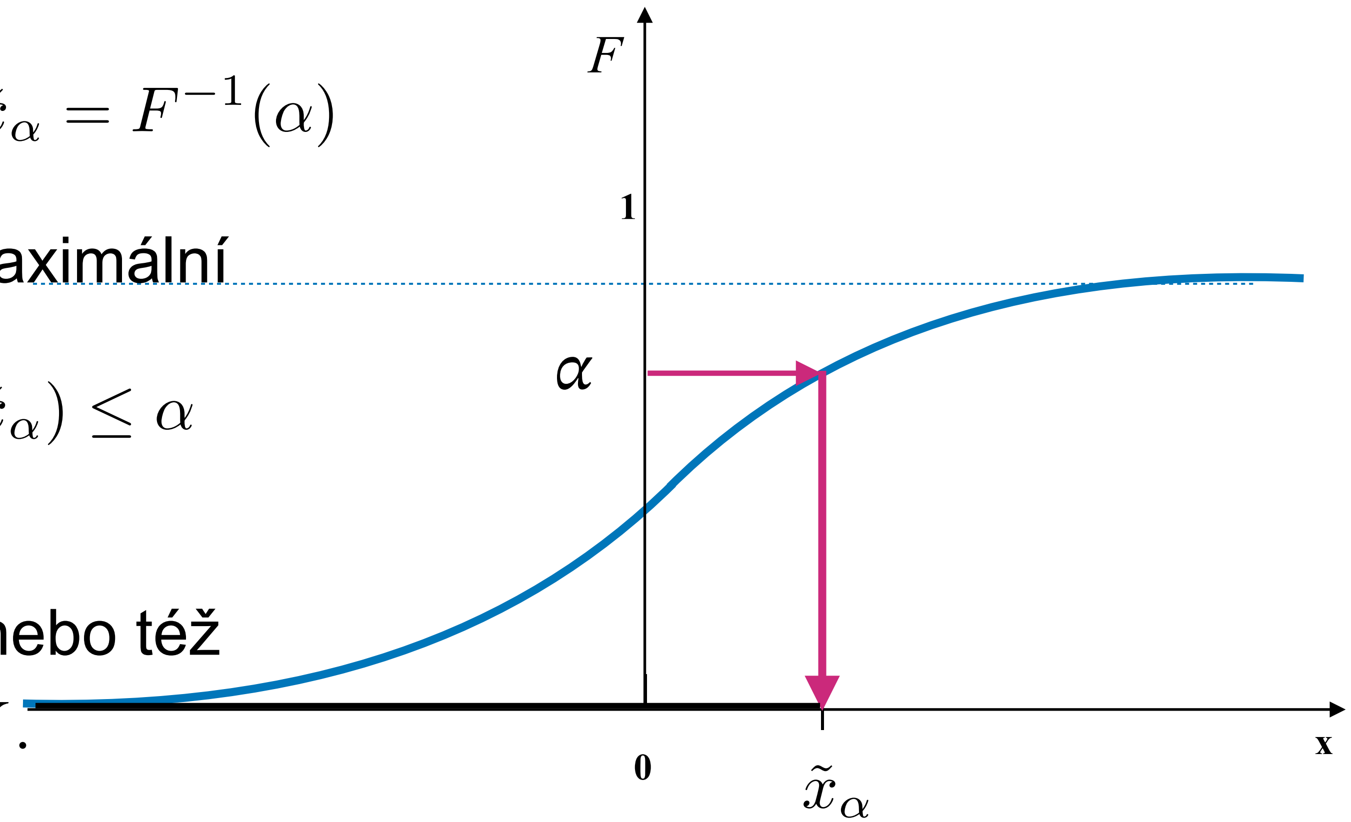
Často se však ptáme: Jakou hodnotu sledovaná náhodná veličina nepřekročí s danou pravděpodobností α?

Příklad: S pravděpodobností α dojde k poruše do doby $\tilde{x}_\alpha$.
Naopak, s pravděpodobností (1-α) součástka (zařízení) vydrží bez poruchy (přežije) po dobu delší než $\tilde{x}_\alpha$.

Charakteristiky náhodné veličiny - Kvantily

Kvantilová funkce: $\tilde{x}_\alpha = F^{-}(\alpha)$

- je-li $F(x)$ spojitá a prostá, potom je $\tilde{x}_\alpha = F^{-1}(\alpha)$
- v ostatních případech je to taková maximální hodnota $\tilde{x}_\alpha$, pro kterou platí: $P(X \leq \tilde{x}_\alpha) \geq \alpha$ a zároveň $P(X < \tilde{x}_\alpha) \leq \alpha$ (tzv. pseudoinverze).
- hodnotu $\tilde{x}_\alpha$ nazýváme α-kvantilem nebo též 100α%-kvantilem náhodné veličiny X.



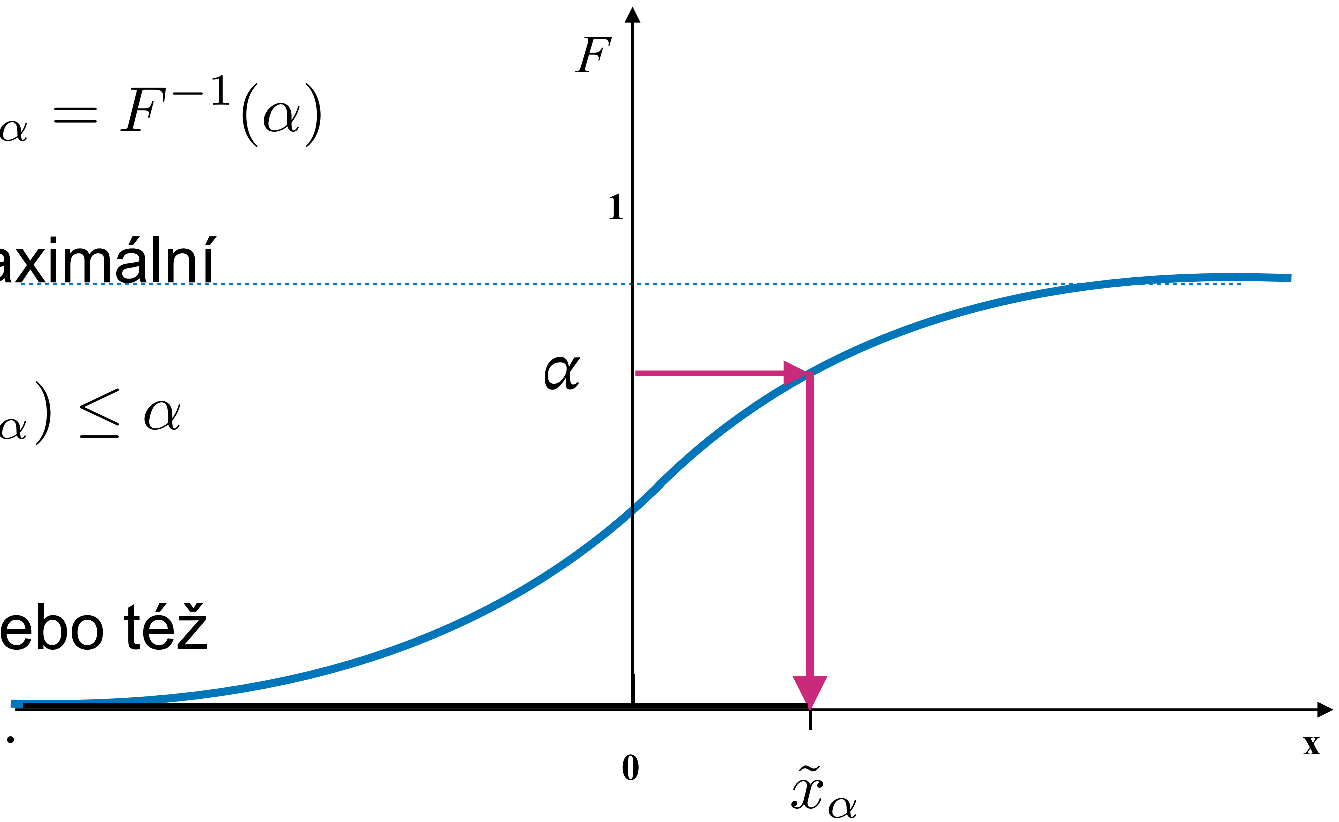
Často se však ptáme: Jakou hodnotu sledovaná náhodná veličina nepřekročí s danou pravděpodobností α?

Příklad: S pravděpodobností α dojde k poruše do doby $\tilde{x}_\alpha$.
Naopak, s pravděpodobností $(1\text{-}\alpha)$ součástka (zařízení) vydrží bez poruchy (přežije) po dobu delší než $\tilde{x}_\alpha$.

Charakteristiky náhodné veličiny - Kvantily

Kvantilová funkce: $\tilde{x}_\alpha = F^{-}(\alpha)$

- je-li $F(x)$ spojitá a prostá, potom je $\tilde{x}_\alpha = F^{-1}(\alpha)$
- v ostatních případech je to taková maximální hodnota $\tilde{x}_\alpha$, pro kterou platí: $P(X \leq \tilde{x}_\alpha) \geq \alpha$ a zároveň $P(X < \tilde{x}_\alpha) \leq \alpha$ (tzv. pseudoinverze).
- hodnotu $\tilde{x}_\alpha$ nazýváme α-kvantilem nebo též 100α%-kvantilem náhodné veličiny X.

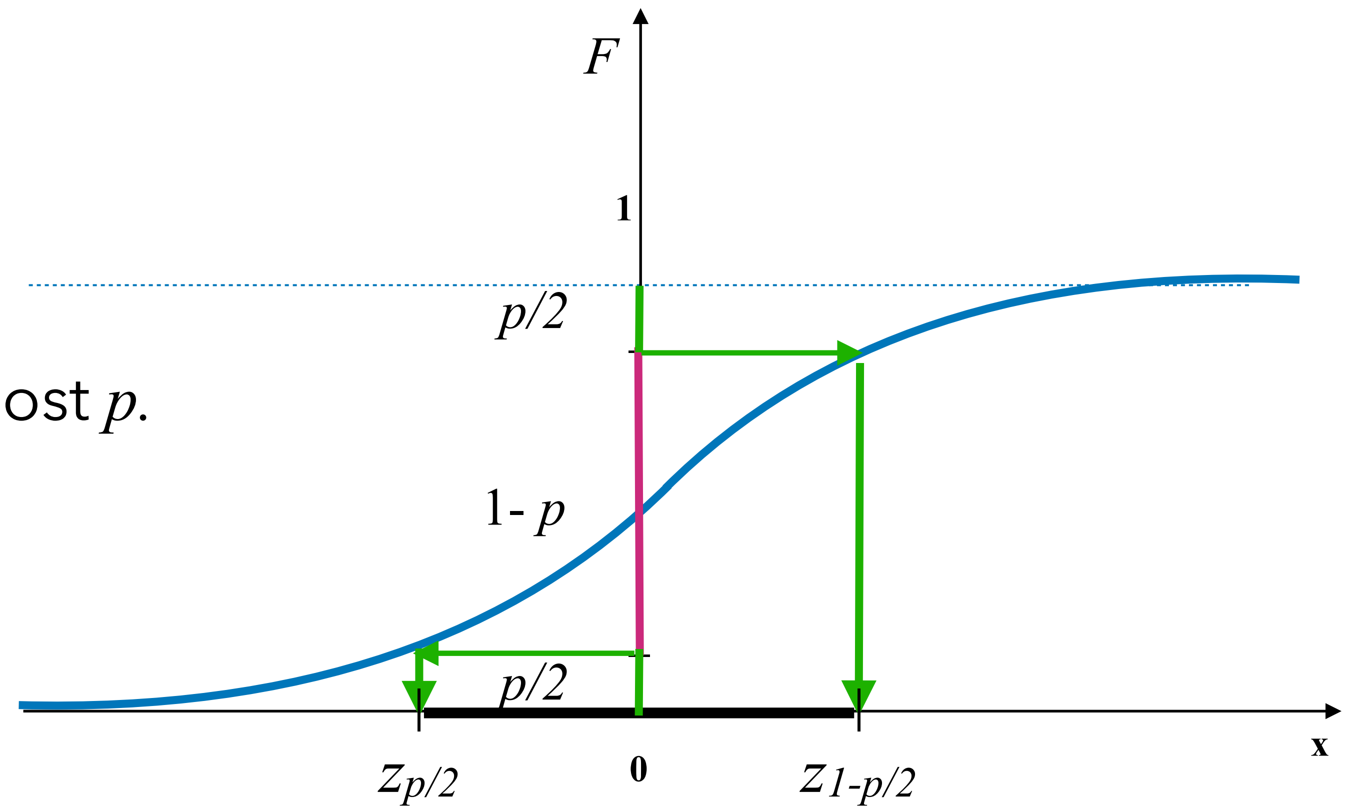


Často se ptáme také: V jakém intervalu lze očekávat hodnoty náhodné veličiny s předem danou pravděpodobností 1-p?

Charakteristiky náhodné veličiny - Kvantily

Obvykle postupujeme takto:

1) Vezmeme zbývající pravděpodobnost p.
2) Rozdělíme ji na dvě poloviny $p/2$:
3) Najdeme odpovídající kvantily
   $z_{p/2}$ a $z_{1-p/2}$
4) Interval $\langle z_{p/2}; z_{1-p/2} \rangle$ je jedním
   z hledaných intervalů.



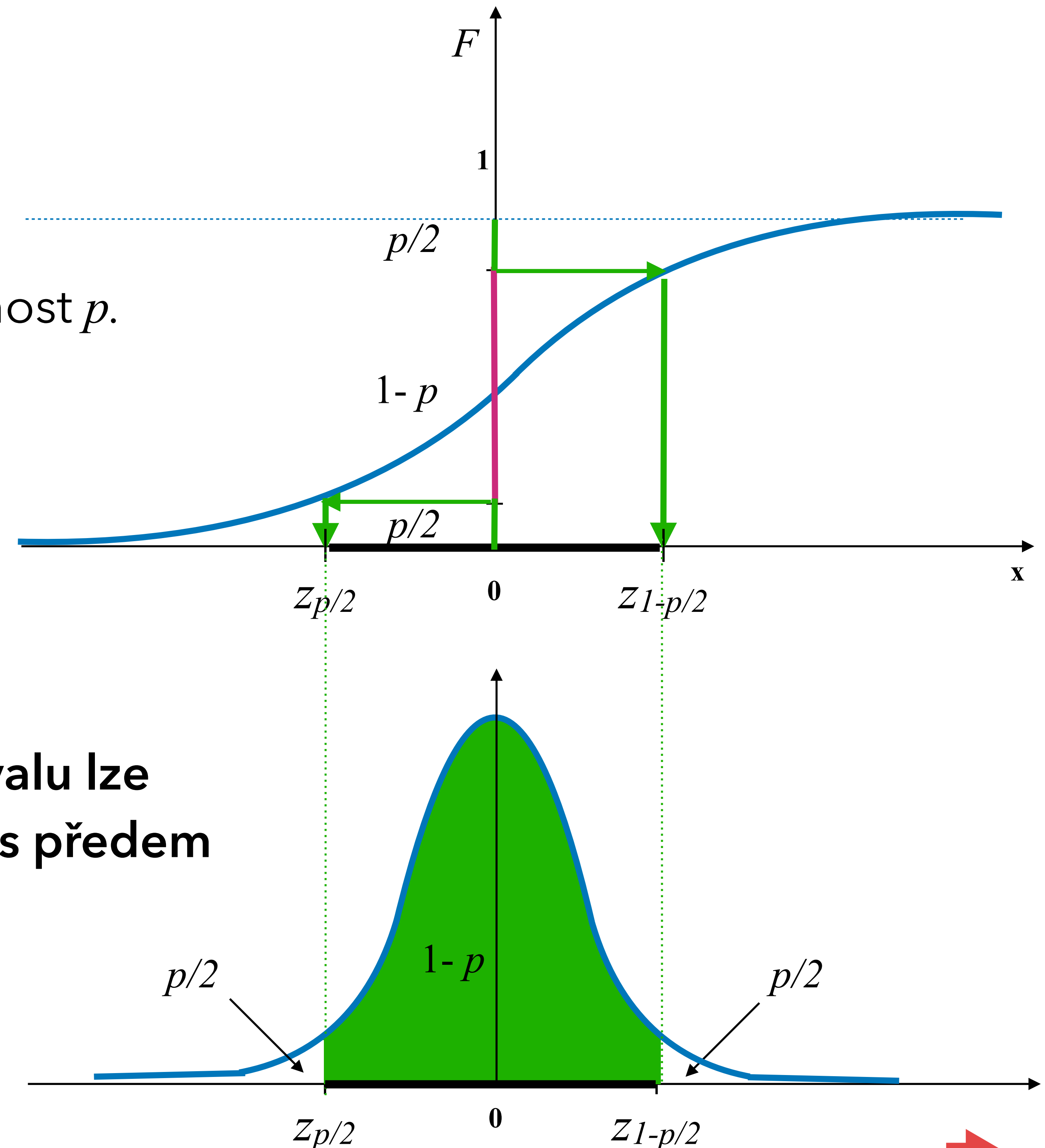
Často se ptáme také: V jakém intervalu lze očekávat hodnoty náhodné veličiny s předem danou pravděpodobností 1-p?

Charakteristiky náhodné veličiny - Kvantily

Obvykle postupujeme takto:

1) Vezmeme zbývající pravděpodobnost p.
2) Rozdělíme ji na dvě poloviny $p/2$:
3) Najdeme odpovídající kvantily

   $z_{p/2}$ a $z_{1-p/2}$
4) Interval $\langle z_{p/2}; z_{1-p/2} \rangle$ je jedním z hledaných intervalů.



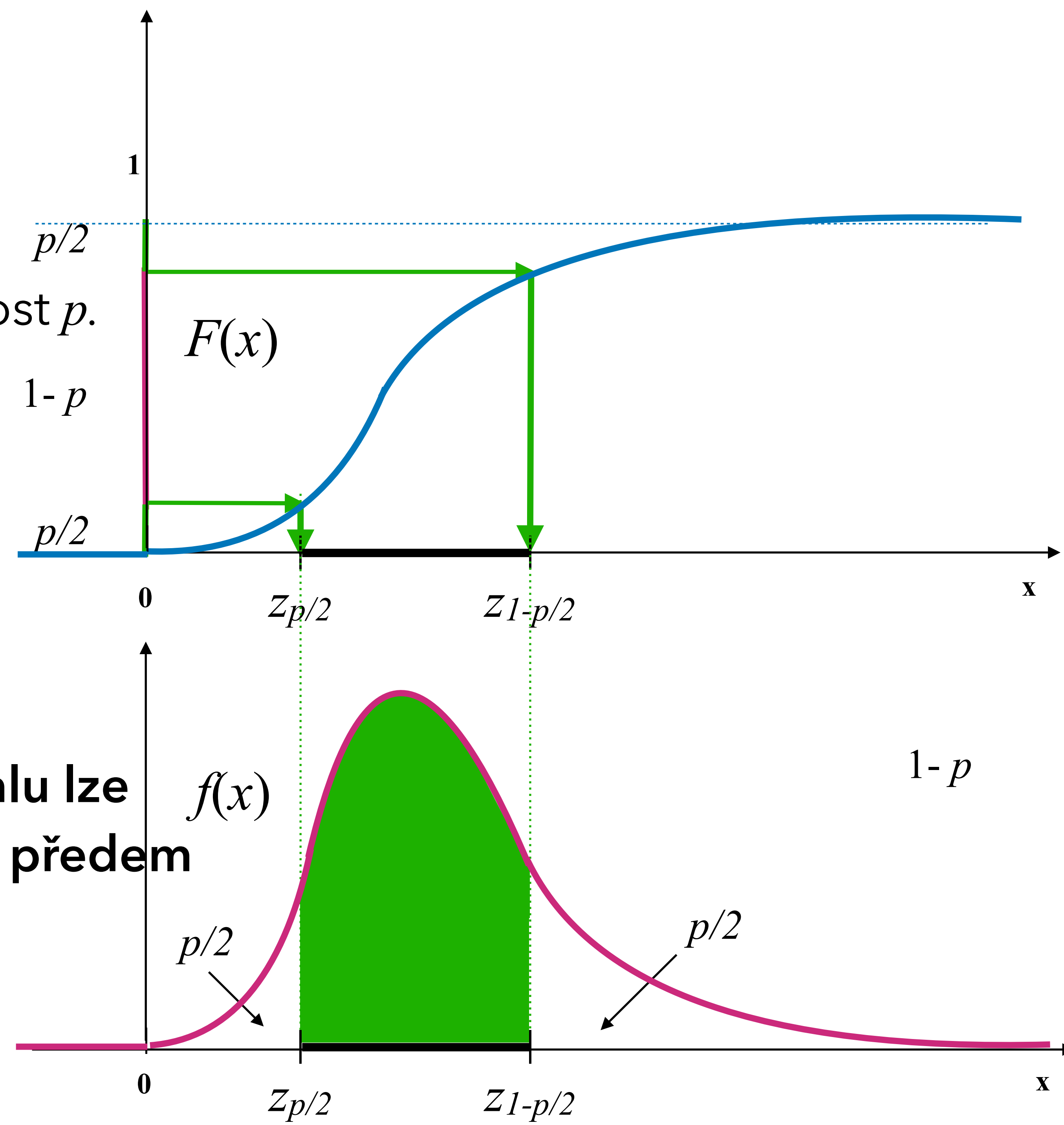
Často se ptáme také: V jakém intervalu lze očekávat hodnoty náhodné veličiny s předem danou pravděpodobností 1-p?

Charakteristiky náhodné veličiny - Kvantily

Obvykle postupujeme takto:

1) Vezmeme zbývající pravděpodobnost p.
2) Rozdělíme ji na dvě poloviny $p/2$:
3) Najdeme odpovídající kvantily

   $z_{p/2}$ a $z_{1-p/2}$
4) Interval $\langle z_{p/2}; z_{1-p/2} \rangle$ je jedním z hledaných intervalů.



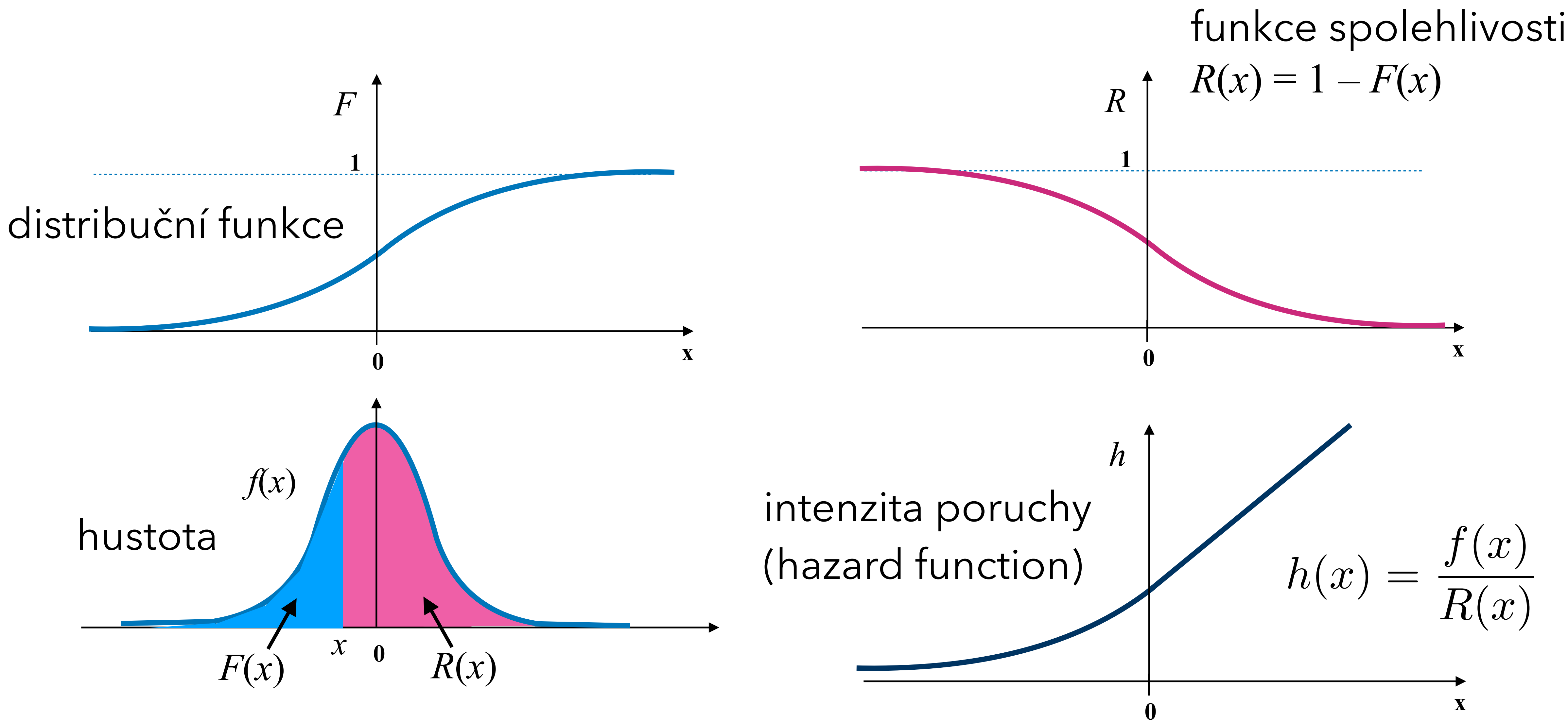
Často se ptáme také: V jakém intervalu lze očekávat hodnoty náhodné veličiny s předem danou pravděpodobností 1-p?

Charakteristiky náhodné veličiny - Kvantily

Často používané kvantily:

- 50%–kvantil = **medián**
- 25%–kvantil = **dolní kvartil**, 75%–kvantil = **horní kvartil**
- 10%–kvantil = **dolní decil**, 90–kvantil = **horní decil**
- (1-α/2)–kvantil = **α–kritická hodnota**

Charakteristiky náhodné veličiny - další ...



Pravděpodobnost, že se zařízení porouchá v čase $x+\varepsilon$, když víme, že se do doby ε neporouchalo (pro hodně malá ε) je rovna přibližně $\varepsilon . h(x)$.

Charakteristiky náhodné veličiny - další ...

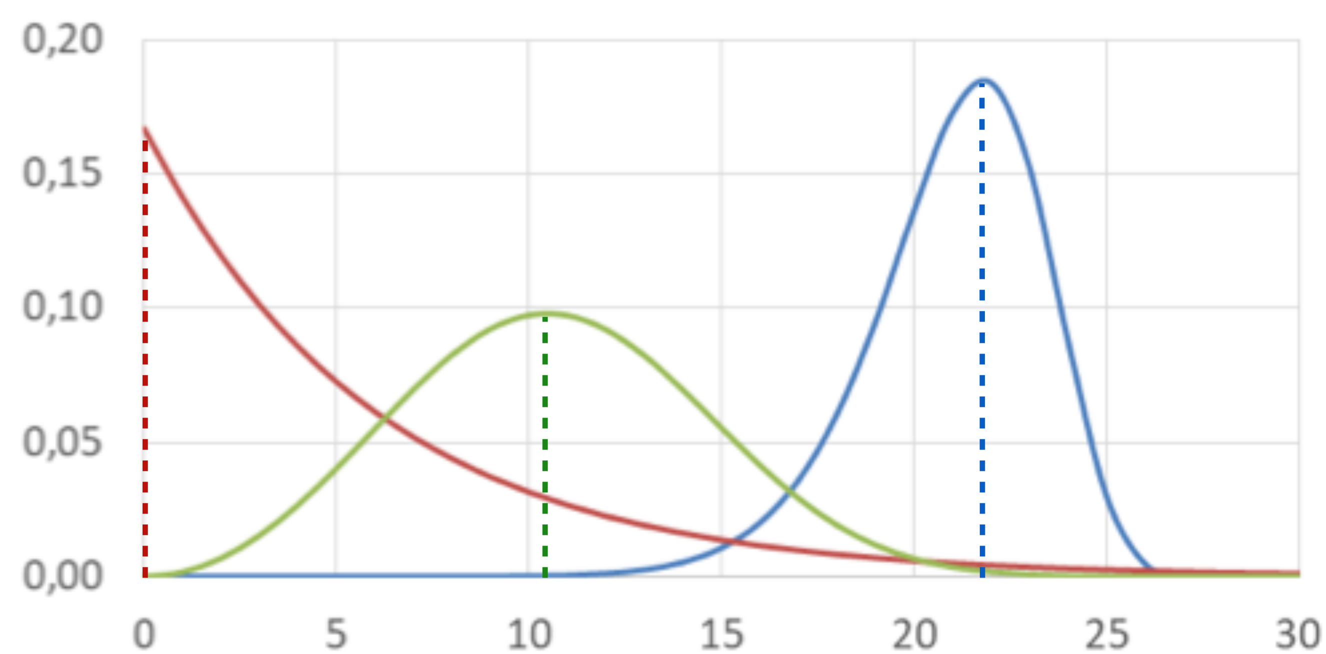
Patří sem například :

- modus (nejpravděpodobnější hodnota)
- rozpětí $(\tilde{x}_1 - \tilde{x}_0)$,
- mezikvartilová odchylka $(\tilde{x}_{0,75} - \tilde{x}_{0,25})$
- a další ….

v diskrétním případě: $x_{\mathrm{mod}} = \max\{p_i; i = 1, 2, \ldots\}$

ve spojitém případě: $x_{\mathrm{mod}} = \mathrm{argmax}\{f(x);\, x \in \mathbf{R}\}$

$$P(X = x_0) = 0$$
$$P(X \in \langle x_0;\, \mathrm{d}x\,\rangle) \doteq f(x_0)\mathrm{d}x$$



Některé charakteristiky pro některá rozdělení nemusejí existovat.

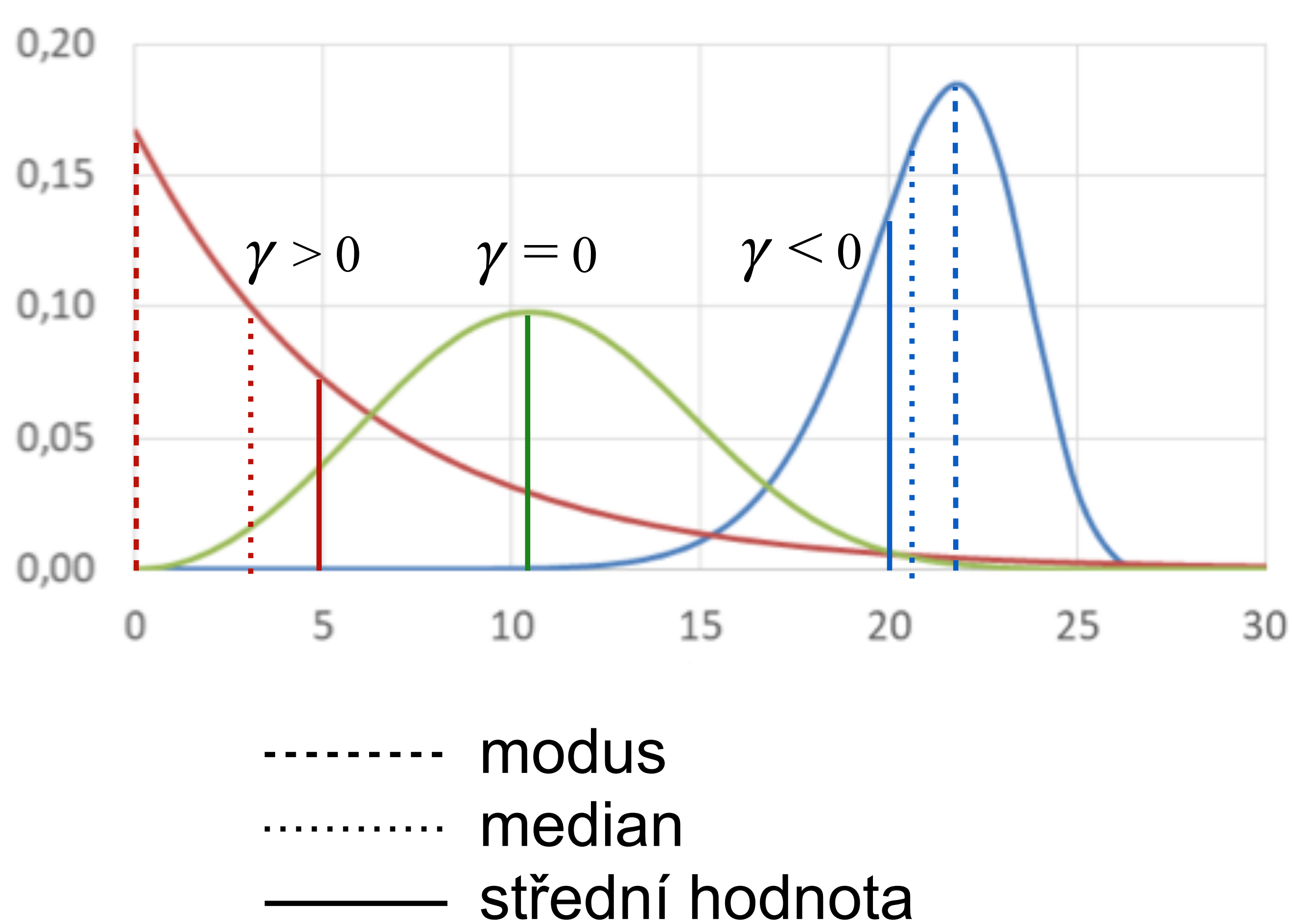
Charakteristiky náhodné veličiny - …

Charakteristiky podle funkce:

- míry polohy - posouvají se s náhodnou veličinou:
  je-li θ mírou polohy veličiny X, potom veličina $X+d$ má míru polohy $\theta+d$
  (patří sem střední hodnota, kvantily, modus, …)

- míry rozptýlenosti (variability) - jsou invariantní vůči posunutí:
  je-li ϑ mírou variability veličiny X, potom je mírou variability i veličiny $X+d$
  (patří sem rozptyl, směrodatná odchylka, různá rozpětí, …)

- míry symetrie (šikmost)



označíme-li šikmost = γ, potom

$\gamma > 0$ – zešikmené zprava (doleva)
$\gamma = 0$ – symetrické
$\gamma < 0$ – zešikmené zleva (doprava)

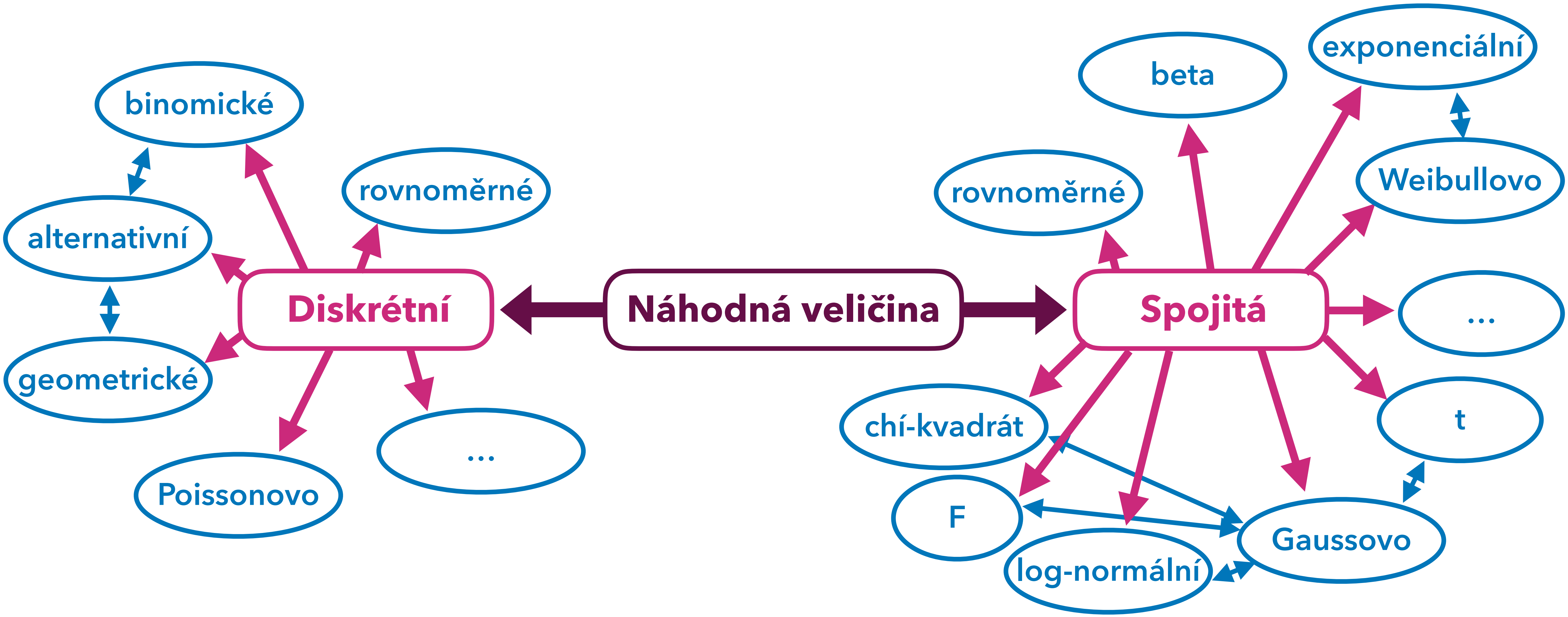
II.1. Základní pojmy a vztahy z teorie pravděpodobnosti

- Náhodná veličina
- Rozdělení pravděpodobnosti
- Charakteristiky rozdělení pravděpodobnosti
- Příklady rozdělení pravděpodobnosti

Rozdělení pravděpodobnosti náhodné veličiny - různé modely

Pravděpodobnostní model = popis experimentu, při němž se realizuje náhodná veličina (jedna či více) s nějakým rozdělením pravděpodobnosti

$\Rightarrow \quad (\Omega, \mathcal{F}, P)$ …. pravděpodobnostní prostor



Rozdělení pravděpodobnosti náhodné veličiny - různé modely

Pravděpodobnostní model = popis experimentu, při němž se realizuje náhodná veličina (jedna či více) s nějakým rozdělením pravděpodobnosti

$\Rightarrow \quad (\Omega, \mathcal{F}, P)$ …. pravděpodobnostní prostor

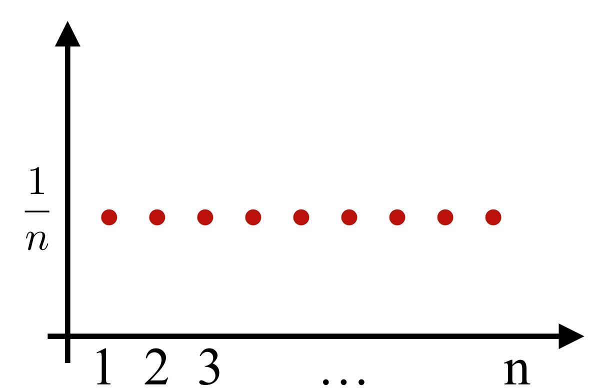
1) Model diskrétního rovnoměrného rozdělení

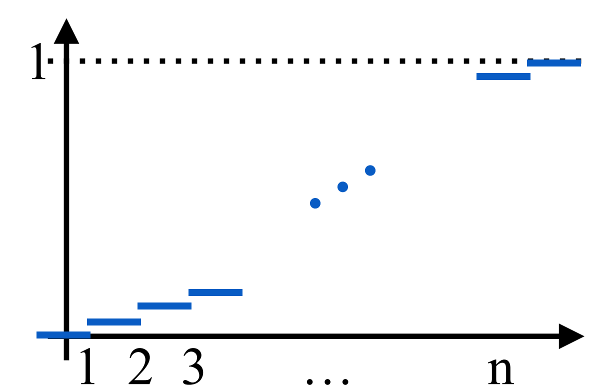
Experiment, při němž může nastat jeden z n možných výsledků. Předpokládáme, že každý z nich je stejně možný, tedy všechny mají stejnou pravděpodobnost.

$\Omega = \{a_1, a_2, \dots, a_n\}, \quad X \in \{x_1, x_2, \dots, x_n\}, \quad P(X = x_i) = 1/n, \quad i = 1, 2, \dots, n.$

$$E(X) = \sum_{i=1}^{n} x_i \frac{1}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}$$

$$Var(X) = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2$$





$X \in \{1, 2, \dots, n\} \Rightarrow E(X) = \frac{n+1}{2} \qquad E(X^2) = \frac{(n+1)(2n+1)}{6}$

$$Var(X) = E(X^2) - (E(X))^2 = \frac{n^2 - 1}{12}$$

Rozdělení pravděpodobnosti náhodné veličiny - různé modely

2) Alternativní model (model alternativního rozdělení)

Experiment, při němž může nastat pouze jeden ze dvou možných výsledků.

$\Omega = \{a_1, a_2\}$, $X \in \{0, 1\}$, $P(X = 1) = p$ => $P(X = 0) = 1 - p$

$$E(X) = 0 \cdot (1-p) + 1 \cdot p = p \qquad E(X^2) = 0^2 \cdot (1-p) + 1^2 \cdot p = p$$

$$Var(X) = E(X^2) - (E(x))^2 = p - p^2 = p(1-p)$$

3) Bernoulliho model (model binomického rozdělení)

Experiment, při němž opakujeme n-krát nezávisle na sobě alternativní pokus za stejných podmínek. Zajímá nás počet „úspěchů“, tedy kolikrát nastal výsledek „1“.

absolutní četnost kladných výsledků = součet pozorování $Y = \sum_{i=1}^{n} X_i$

$$\boxed{P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n.} \qquad Y \sim Bin(n, p)$$

Cvičení: 1) Ukažte, že součet těchto pravděpodobností je roven jedné.
2) Spočtěte střední hodnotu a rozptyl veličiny Y podle definice.

Rozdělení pravděpodobnosti náhodné veličiny - různé modely

3) Bernoulliho model (model binomického rozdělení)

absolutní četnost kladných výsledků = součet pozorování $Y = \sum_{i=1}^{n} X_i$

$$E(Y) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = np \qquad Var(Y) = E(Y - np)^2 = np(1-p)$$

relativní četnost kladných výsledků = aritmetický průměr pozorování $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}np = p$$

$$Var(\bar{X}) = E(\bar{X} - p)^2 = \frac{p(1-p)}{n}$$

Rozdělení pravděpodobnosti náhodné veličiny - různé modely

3) Bernoulliho model (model binomického rozdělení)

Experiment, při němž z množiny N prvků, mezi nimiž je M prvků s určitou vlastností, vybíráme náhodně postupně n prvků tak, že po každém výběru vybraný prvek vrátíme zpět mezi ostatní. Zajímá nás kolikrát vybereme prvek s danou vlastností.

P(vybereme prvek s danou vlastností) $= \dfrac{M}{N}$, $\quad Y =$ počet vybraných prvků s danou vlastností

$$P(Y = k) = \binom{N}{n} \left(\frac{M}{N}\right)^k \left(1 - \frac{M}{N}\right)^{n-k} \qquad\qquad Y \sim Bin\Big(n, \frac{M}{N}\Big)$$

počet k-tic
v n prvcích

k-krát vybereme prvek
s pravděpodobností $\frac{M}{N}$

(n-k)-krát vybereme prvek
s pravděpodobností (1- $\frac{M}{N}$)

Rozdělení pravděpodobnosti náhodné veličiny - různé modely

3) Bernoulliho model (model binomického rozdělení)

Experiment, při němž z množiny N prvků, mezi nimiž je M prvků s určitou vlastností, vybíráme náhodně postupně n prvků tak, že po každém výběru vybraný prvek vrátíme zpět mezi ostatní. Zajímá nás kolikrát vybereme prvek s danou vlastností.

$$\mathrm{P}(\text{vybereme prvek s danou vlastností}) = \frac{M}{N}, \quad Y = \text{počet vybraných prvků s danou vlastností}$$

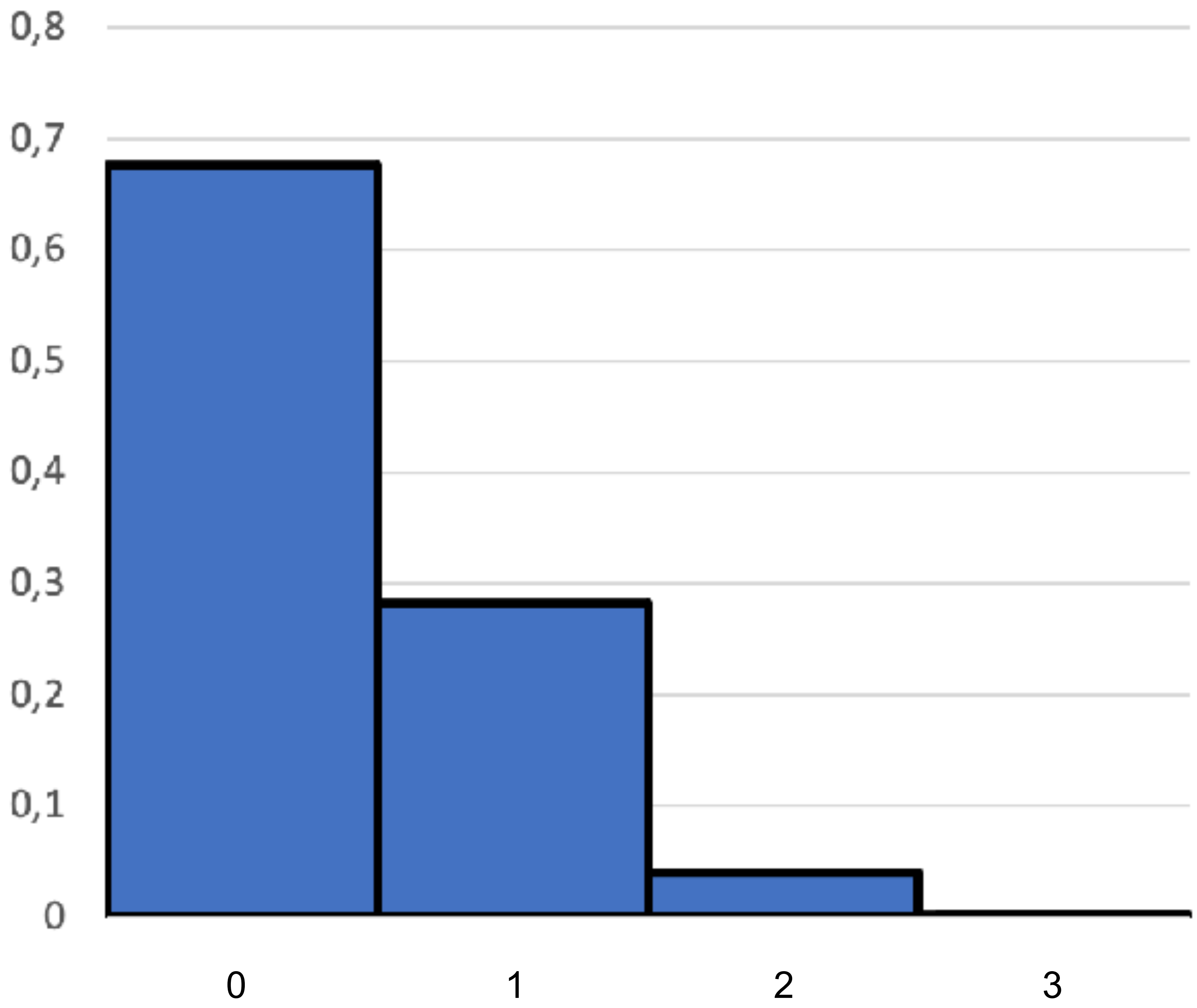
$$P(Y = k) = \binom{N}{n} \left(\frac{M}{N}\right)^k \left(1 - \frac{M}{N}\right)^{n-k} \qquad Y \sim Bin\left(n, \frac{M}{N}\right)$$

Příklad 1: Házení kostkou: házíme třemi hracími kostkami současně (nebo jednou třikrát po sobě). Jaká je pravděpodobnost, že padnou alespoň dvě šestky?

Příklad 2: Kontrola jakosti: Z výrobní linky odebíráme nezávisle na sobě 10 výrobků. Víme, že v produkci jsou 3% vadných. Jaká je pravděpodobnost, že při výběru vybereme alespoň 1 zmetek?

Příklad 3: Losování zaměstnance: každý den v týdnu losujeme jednoho z 15ti zaměstnanců, který provede odpolední úklid. Mezi zaměstnanci 10 žen. Jaká je pravděpodobnost, že v týdnu vybereme samé ženy?

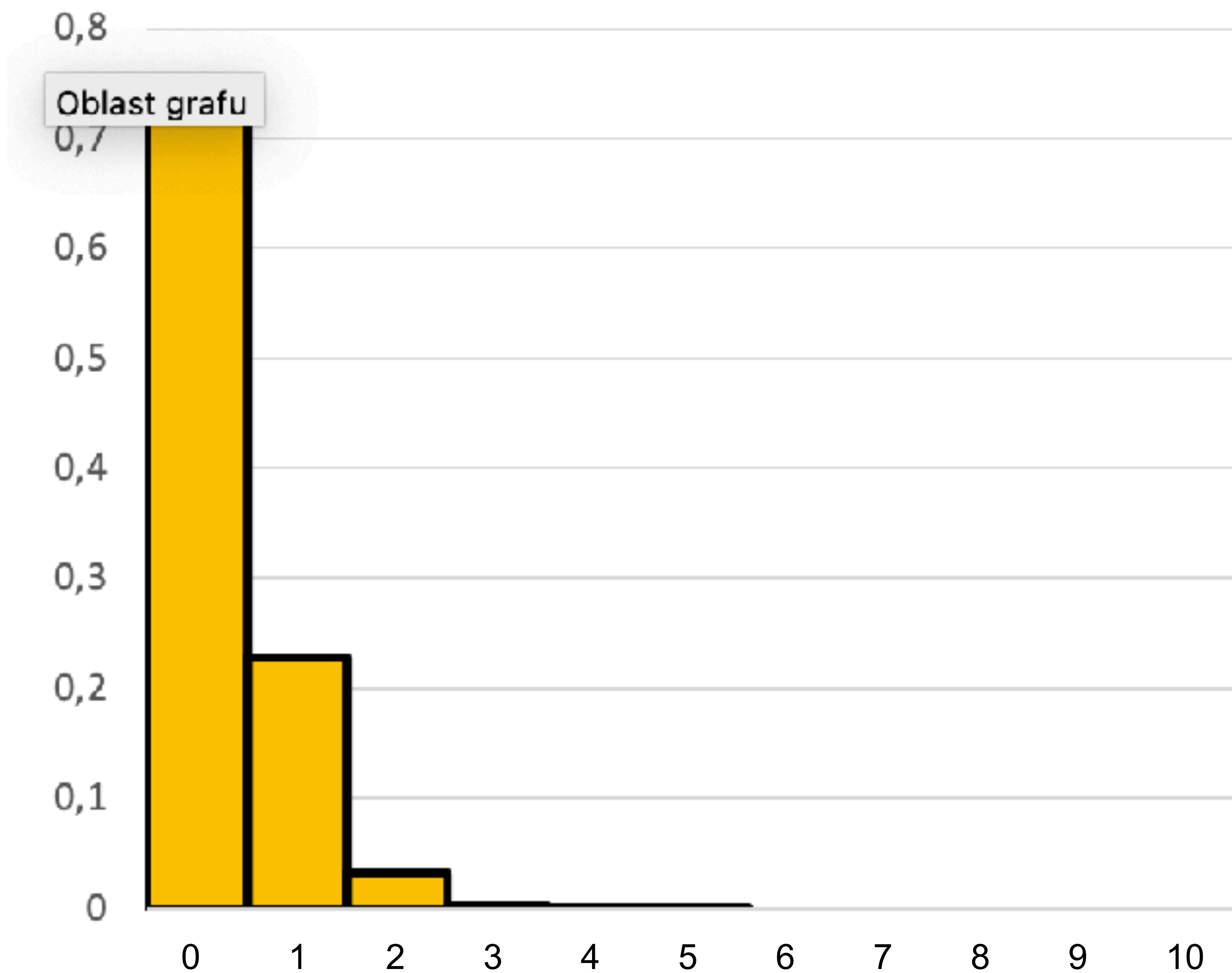
Příklad 1: Házení kostkou: házíme třemi hracími kostkami současně (nebo jednou třikrát po sobě). Jaká je pravděpodobnost, že padnou alespoň dvě šestky?



X = počet šestek

k	P(X=k)
0	0,6757983
1	0,2828923
2	0,0394733
3	0,001836

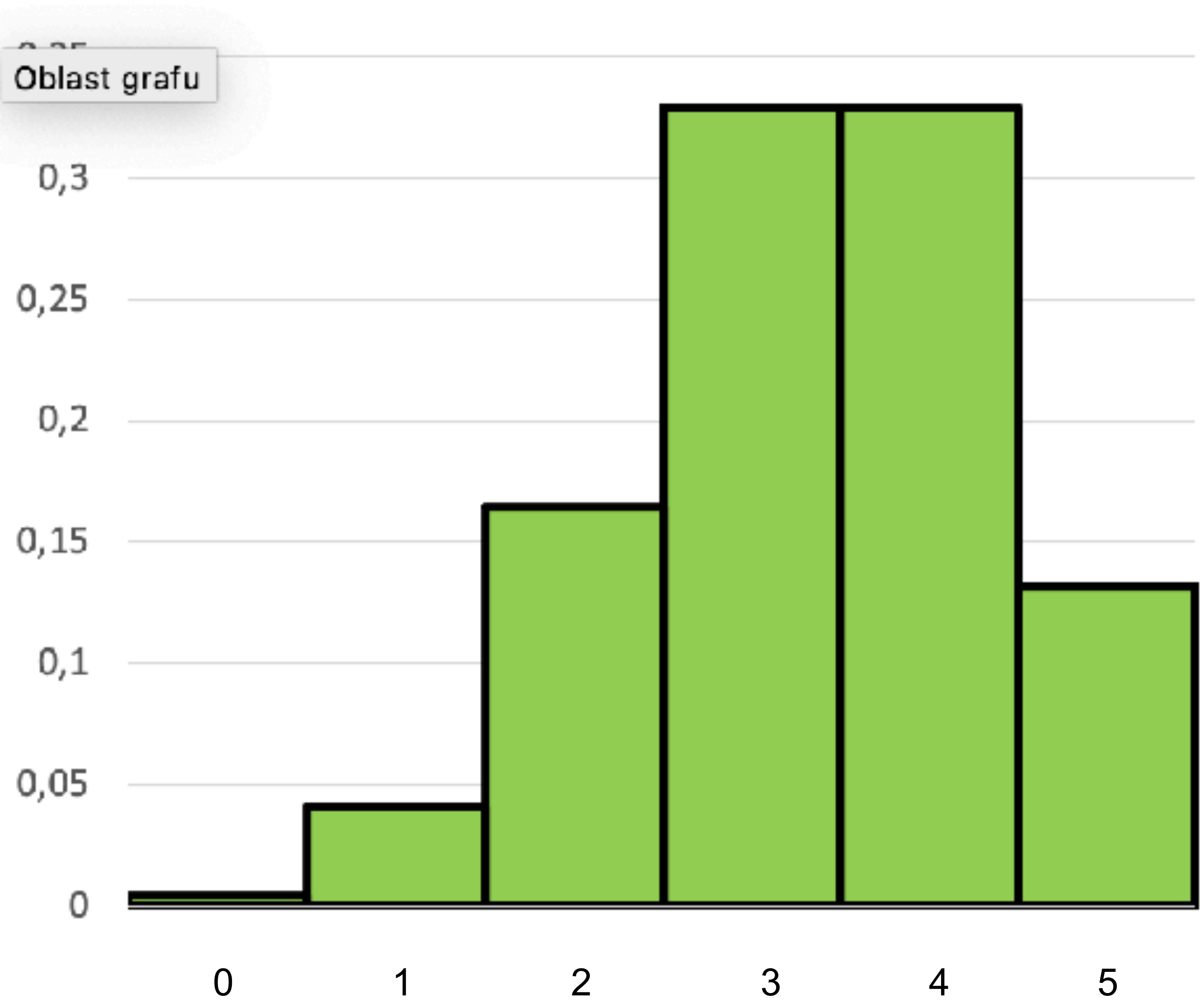
Příklad 2: Kontrola jakosti: Z výrobní linky odebíráme nezávisle na sobě 10 výrobků. Víme, že v produkci jsou 3% vadných. Jaká je pravděpodobnost, že při výběru vybereme alespoň 1 zmetek?



X = počet vybraných zmetků

k	P(X=k)
0	0,7374241
1	0,2280693
2	0,0317416
3	0,0026179
4	0,0001417
5	5,259E-06
6	1,355E-07
7	2,395E-09
8	2,778E-11
9	1,909E-13
10	5,905E-16

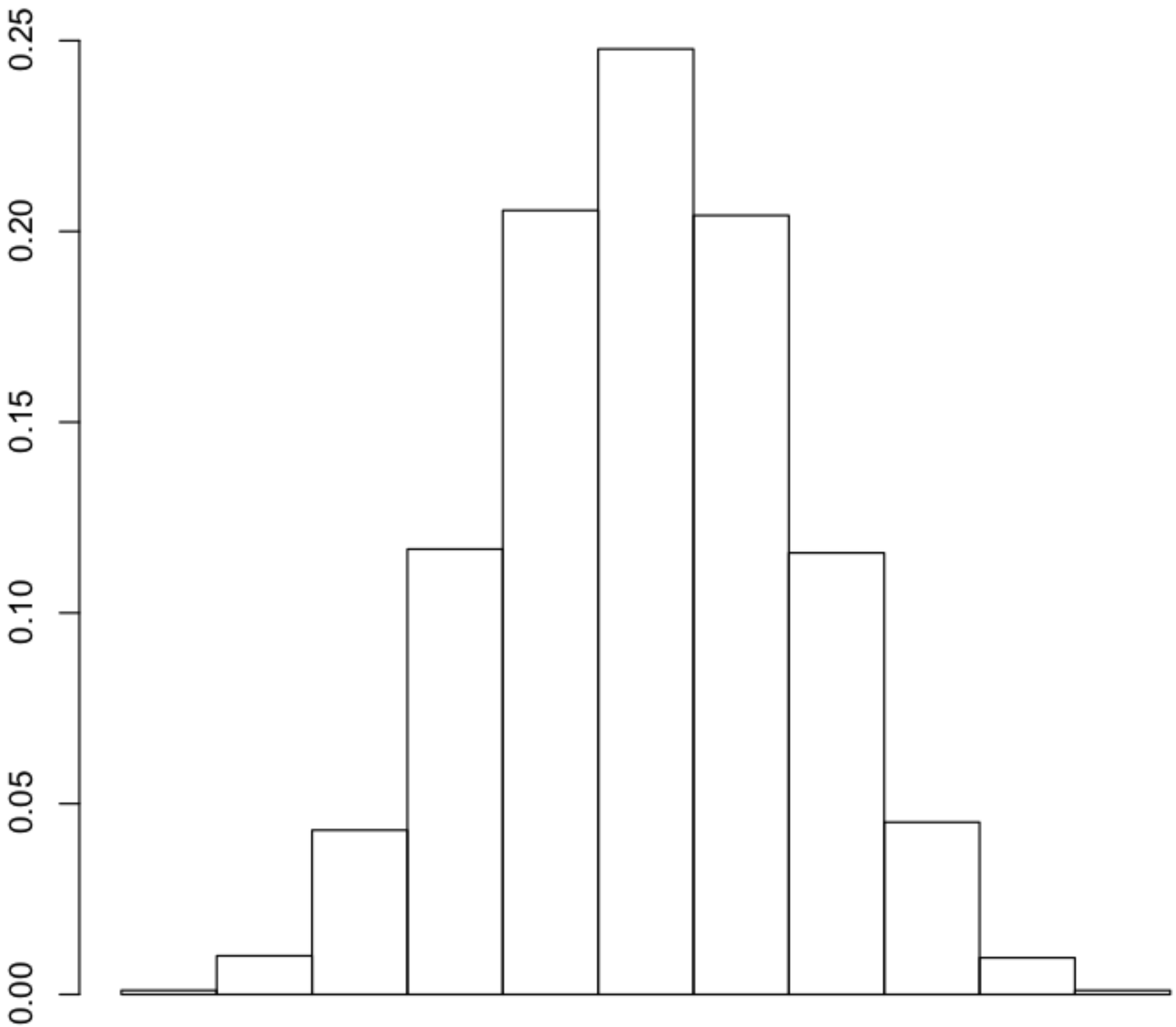
Příklad 3: Losování zaměstnance: každý den v týdnu losujeme jednoho z 15ti zaměstnanců, který provede odpolední úklid. Mezi zaměstnanci 10 žen. Jaká je pravděpodobnost, že v týdnu vybereme samé ženy?



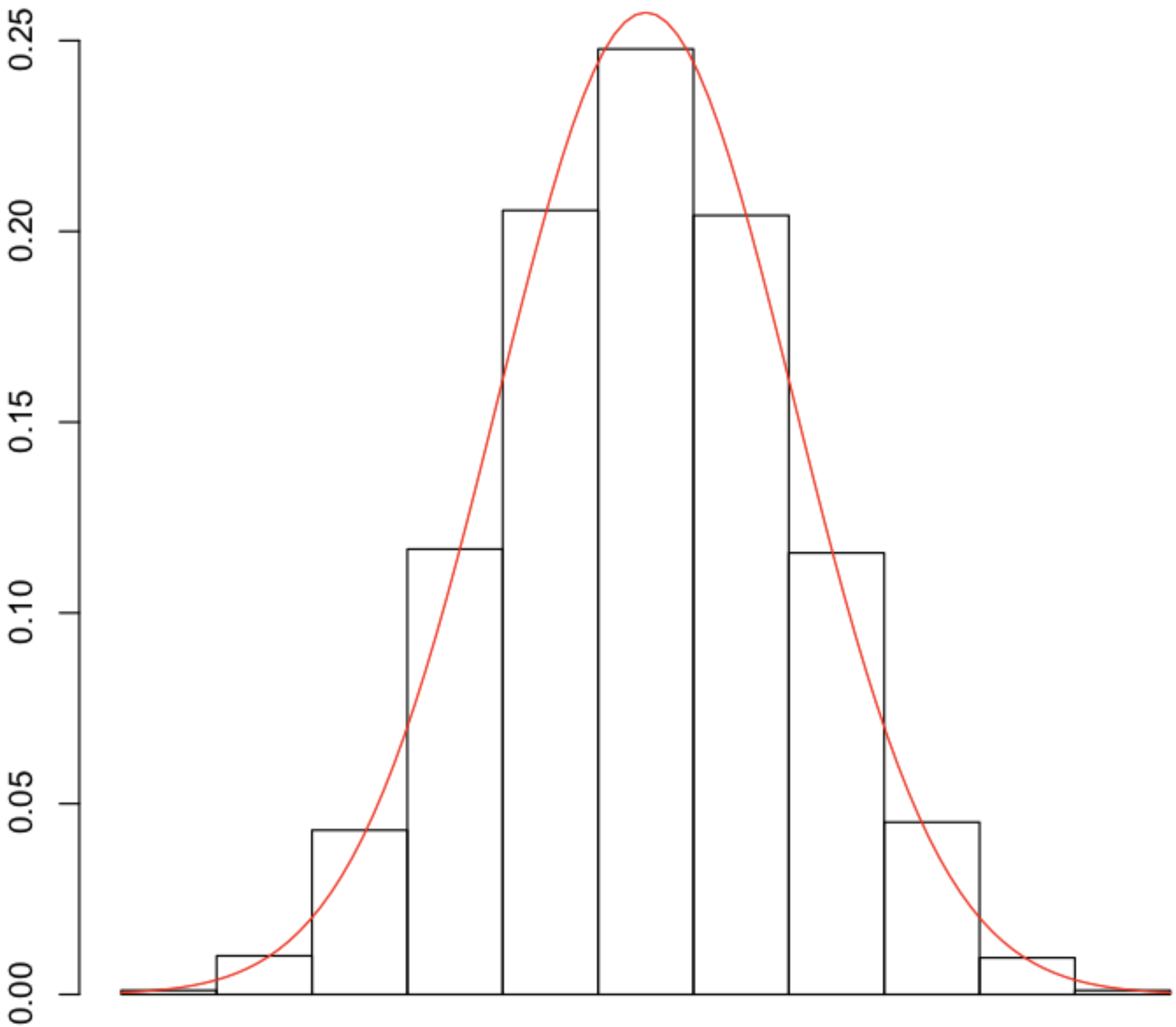
X = počet vybraných žen

k	P(X=k)
0	0,0041152
1	0,0411523
2	0,1646091
3	0,3292181
4	0,3292181
5	0,1316872

Binomické rozdělení lze pro velká n a relativně malá k aproximovat spojitým normálním rozdělením



Binomické rozdělení lze pro velká n a relativně malá k aproximovat spojitým normálním rozdělením



Rozdělení pravděpodobnosti náhodné veličiny - různé modely

4) Hypergeometrický model (model výběru bez vracení)

Experiment, při němž z množiny N prvků, mezi nimiž je M prvků s určitou vlastností, vybíráme náhodně postupně n prvků tak, že po každém výběru vybraný prvek nevracíme zpět mezi ostatní. Zajímá nás kolik prvků s danou vlastností vybereme.

k prvků s danou vlastností z celkového počtu M můžeme vybrat celkem $\binom{M}{k}$ způsoby

zbylých n-k prvků bez této vlastnost můžeme vybrat celkem $\binom{N-M}{n-k}$ způsoby

celkový počet možných n výběrů z N prvků je $\binom{N}{n}$

Tedy: $$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

$$N=1,2,\ldots,\quad M\le N,\quad n\le N,$$
$$\max\{\,0,\,n+M-N\,\}\le k\le\min\{\,n,\,M\,\}$$

$$E(X)=n\frac{M}{N}$$

$$Var(X)=\frac{nM(N-n)(N-M)}{N^2(N-1)}$$

Hypergeometrické rozdělení
s parametry N, M, n.

Rozdělení pravděpodobnosti náhodné veličiny - různé modely

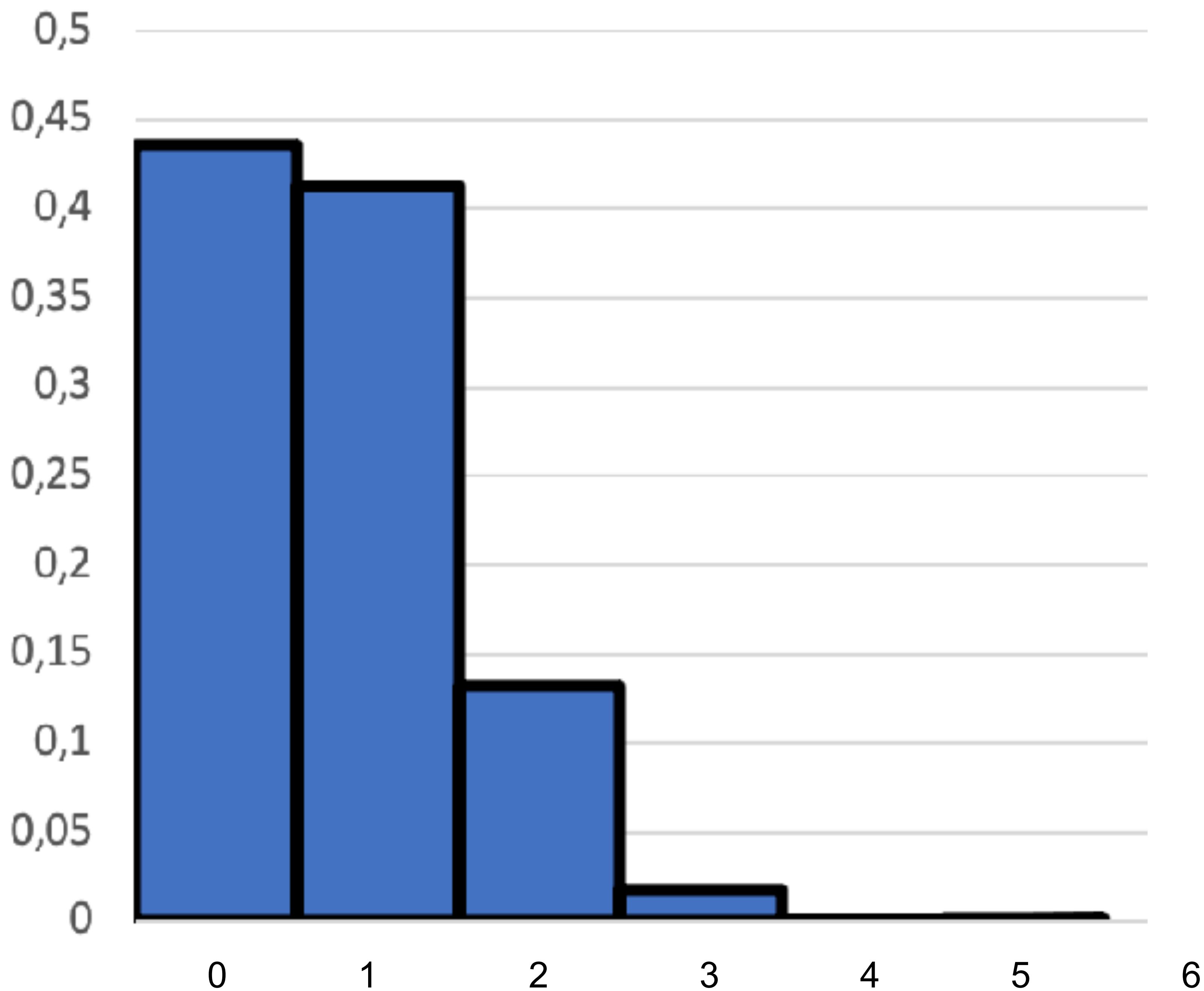
4) Hypergeometrický model (model výběru bez vracení)

Příklad 4: Sportka: 49 čísel, ze kterých 6 vyhrává (jsou vytaženy). Jaká je pravděpodobnost, že při výběru 6ti čísel vybereme 4 z tažených?

Příklad 5: Kontrola jakosti: 1000 výrobků, mezi nimi jsou 3% vadných. Jaká je pravděpodobnost, že při výběru 10 výrobků vybereme alespoň 1 zmetek?

Příklad 6: Výběr uchazečů o práci: z 15ti uchazečů o zaměstnání, mezi kterými je 10 žen, vybíráme anonymně podle výsledku testu 5 osob.
Jaká je pravděpodobnost, že to budou samé ženy?

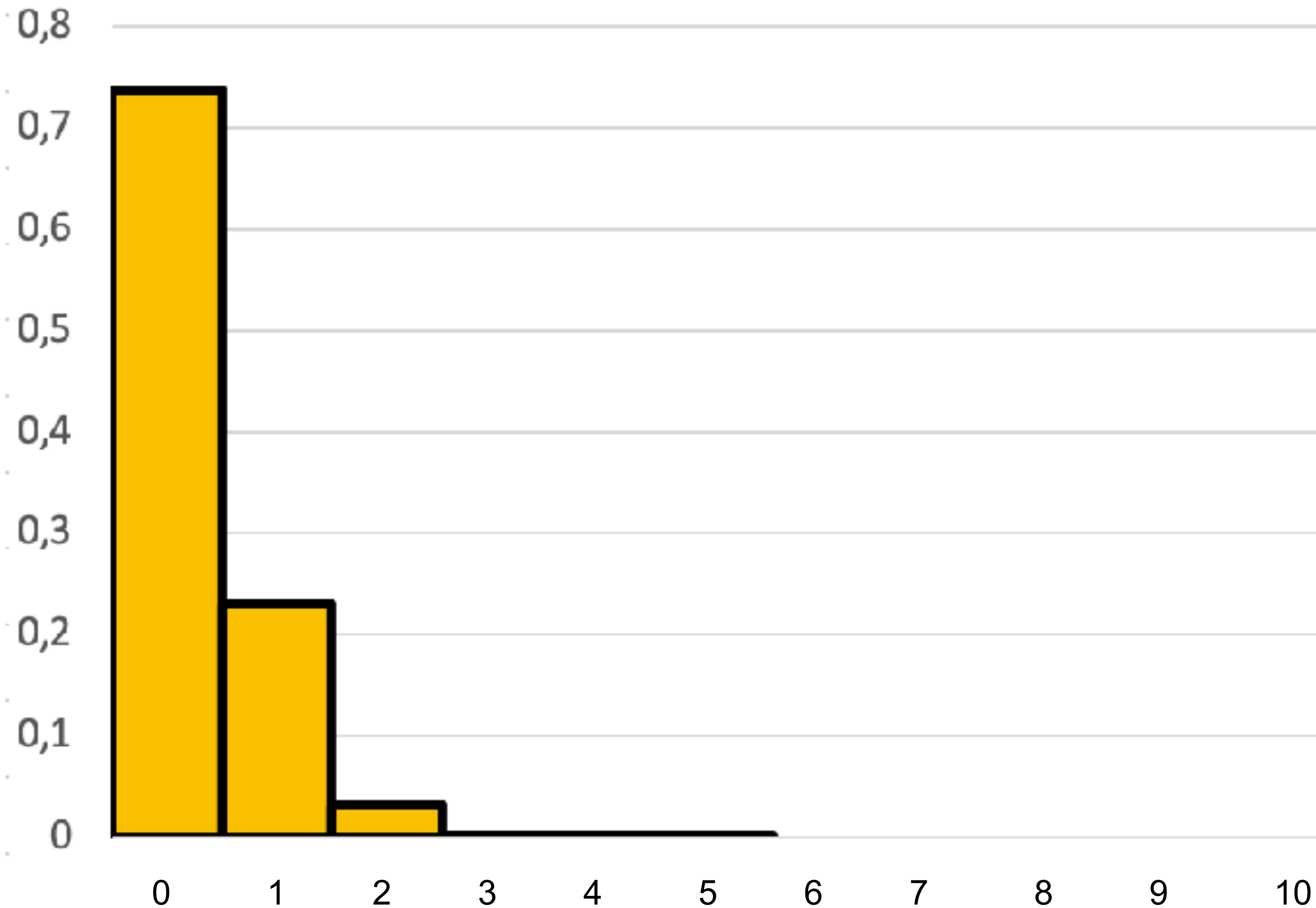
Příklad 4: Sportka: 49 čísel, ze kterých 6 vyhrává (jsou vytaženy). Jaká je pravděpodobnost, že při výběru 6ti čísel vybereme 4 z tažených?



X = počet uhodnutých tažených čísel

k	P(X=k)
0	0,43596498
1	0,41301945
2	0,13237803
3	0,0176504
4	0,00096862
5	1,845E-05
6	7,1511E-08

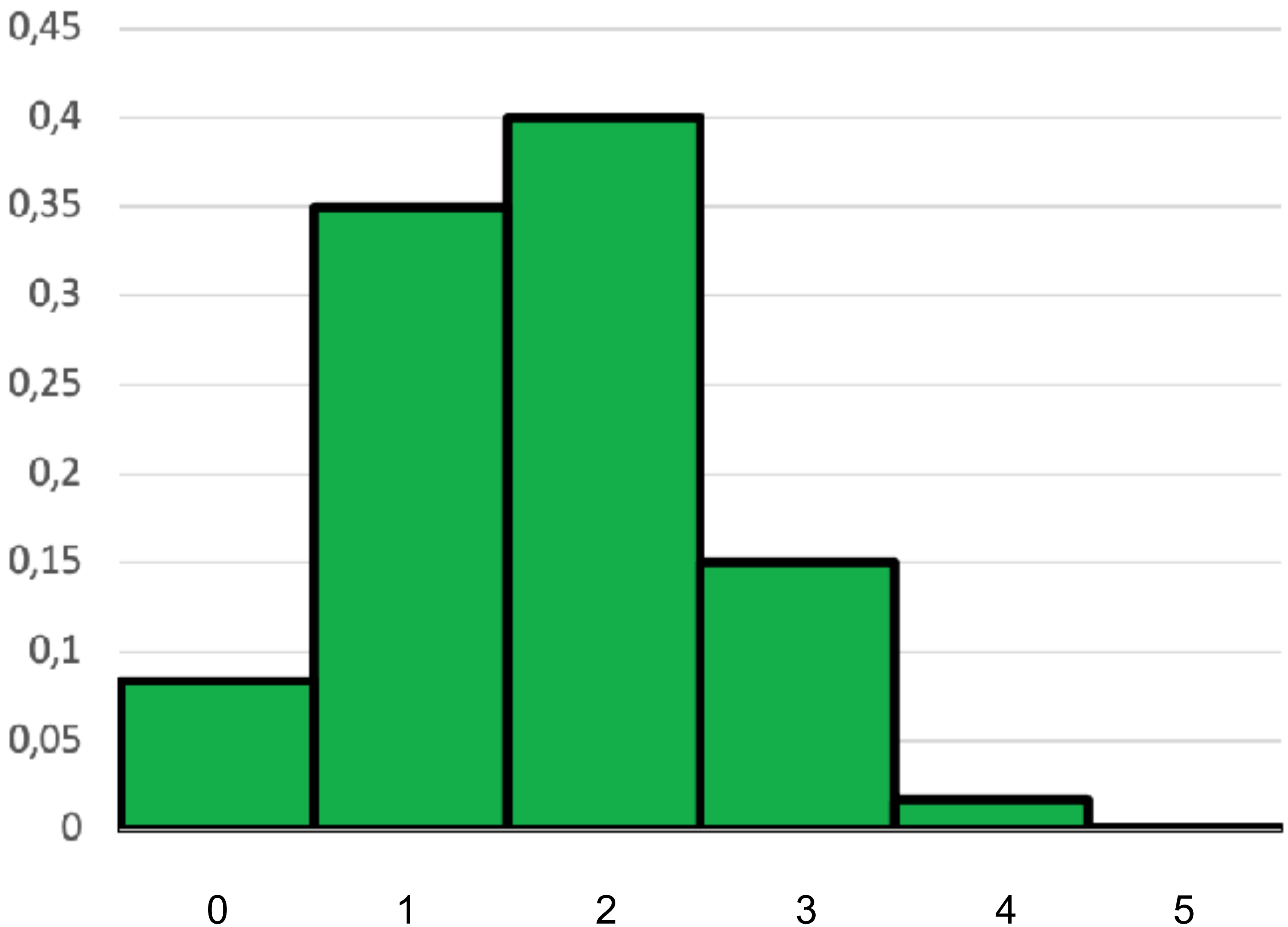
Příklad 5: Kontrola jakosti: 1000 výrobků, mezi nimi jsou 3% vadných. Jaká je pravděpodobnost, že při výběru 10 výrobků vybereme alespoň 1 zmetek?



X = počet vybraných zmetků

k	P(X=k)
0	0,73639189
1	0,229883
2	0,03118475
3	0,00241792
4	0,00011851
5	3,8317E-06
6	8,2637E-08
7	1,172E-09
8	1,0443E-11
9	5,2686E-14
10	1,1406E-16

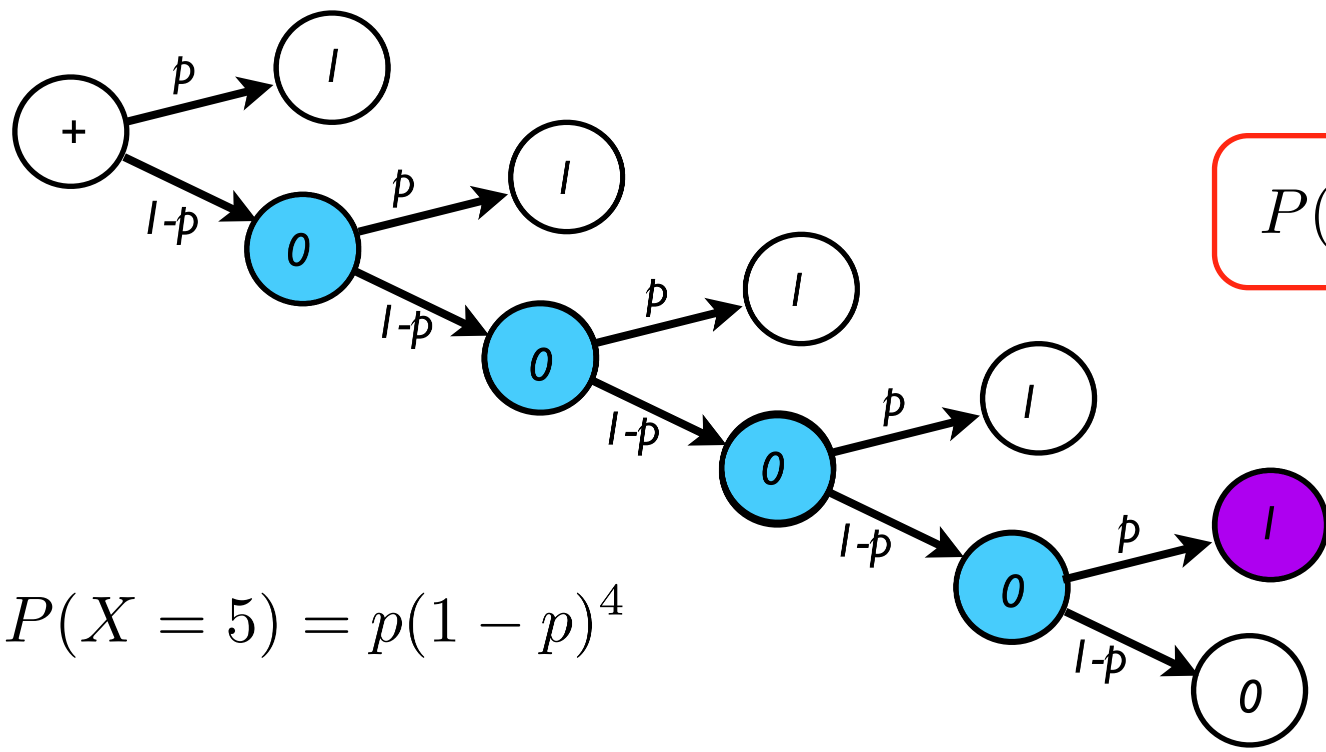
Příklad 6: Výběr uchazečů o práci: z 15ti uchazečů o zaměstnání, mezi kterými je 10 žen, vybíráme anonymně podle výsledku testu 5 osob. Jaká je pravděpodobnost, že to budou samé ženy?



X = počet mužů ve výběru

k	P(X=k)
0	0,08391608
1	0,34965035
2	0,3996004
3	0,14985015
4	0,01665002
5	0,000333

5) Geometrický model (diskrétní doba života)



$$P(X = 5) = p(1-p)^4$$

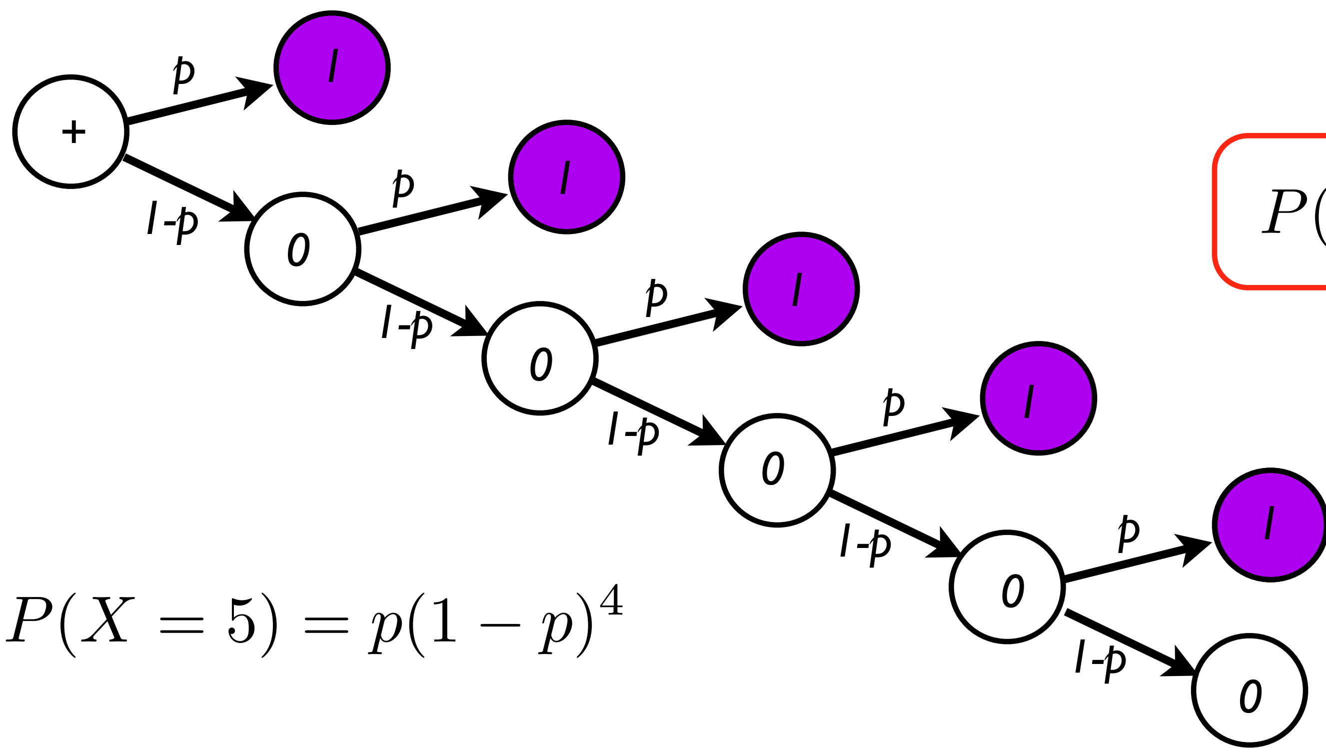
$$\boxed{P(X = k) = p(1-p)^{k-1}}$$

$$k = 1, 2, \ldots$$

X je počet kroků, které je třeba učinit, aby nastal první výskyt sledovaného jevu

$$E(X) = \frac{1}{p} \qquad Var(X) = \frac{1-p}{p^2}$$

5) Geometrický model (diskrétní doba života)



$$P(X = k) = p(1-p)^{k-1}$$

$$k = 1, 2, \ldots$$

$$P(X = 5) = p(1-p)^4$$

X je počet kroků, které je třeba učinit, aby nastal první výskyt sledovaného jevu

$$E(X) = \frac{1}{p} \qquad Var(X) = \frac{1-p}{p^2}$$

Y je počet kroků, které předcházejí prvnímu výskytu sledovaného jevu

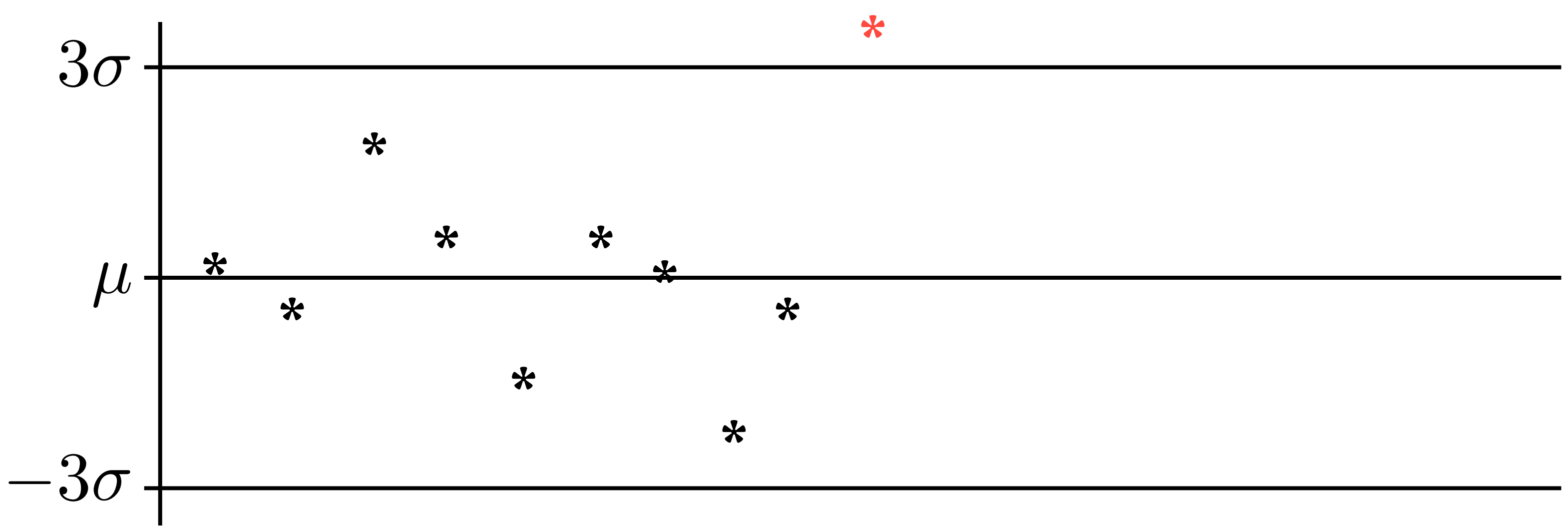
$$P(Y = k) = p(1-p)^k$$

$$k = 0, 1, \ldots$$

$$E(Y) = \frac{1-p}{p} \qquad Var(Y) = \frac{1-p}{p^2}$$

Je-li sledovaný jev porucha, potom se Y nazývá “diskrétní doba života”

5) Geometrický model (diskrétní doba života)



$$P(|X-\mu| \leq 3\sigma) = \Phi(3) - \Phi(-3) = 2\Phi(3) - 1 = 0,9973$$

$$P(|X-\mu| \geq 3\sigma) = 0,0027$$

N = počet inspekcí před signálem

$p = 0,0027$ $\qquad$ $E(N) = \frac{1}{p} = \frac{1}{0,0027} = 370$

Počet inspekcí před prvním falešným signálem (ARL = Average Run Length)