VI. Spojité pravděpodobnostní modely

Náhodná veličina - opakování

- **Statistický soubor** tvoří statistické jednotky (např. zákazníci)
- Sledujeme (kvantitativní) hodnotu statistického znaku (např. cena nákupů)
- každá jednotka = jedna hodnota (číslo) statistického znaku (např. 453,55 Kč)

Náhodná veličina - opakování

- **Náhodná veličina** = soubor všech hodnot znaku + rozdělení pr-sti hodnot

  - některé hodnoty se nabývají častěji než jiné → mají větší pravděpodobnost výskytu

  - hodnoty znaku statistických jednotek se „generují“ podle pravděpodobnostního rozdělení daného:

  (1) distribuční funkcí (2) hustotou pr-sti

Spojitá náhodná veličina – příklady 1

Spojitá náhodná veličina = **NV**, kde možnými hodnotami jsou všechna reálná čísla z daného intervalu (omezeného nebo neomezeného), např.:

1. Čísla generovaná v Excelu pomocí funkce NÁHČÍSLO (v anglické verzi RAND) je náhodná veličina X, která může teoreticky nabývat jakékoliv hodnoty mezi 0 a 1

2. Výsledky měření důležitých rozměrů nebo hmotnosti produktů je náhodná veličina X, která může teoreticky nabývat jakékoliv hodnoty $x > 0$

Spojitá náhodná veličina – příklady 2

Spojitá náhodná veličina = NV, kde možnými hodnotami jsou všechna reálná čísla z daného intervalu (omezeného nebo neomezeného), např.:

3. Doba strávená zákazníkem ve frontě u přepážky v bance je náhodná veličina X, která může teoreticky nabývat jakékoliv hodnoty $x \geq 0$

4. Cena akcií je náhodná veličina X, která může teoreticky nabývat jakékoliv hodnoty $x \geq 0$

akcie ČEZ

Rozdělení náhodné veličiny (NV): distribuční funkce a funkce hustoty

- *distribuční* ***funkce*** **(DF)** NV X - každému reálnému číslu přiřazuje pravděpodobnost, že náhodná veličina X nabude hodnoty ne větší, než toto číslo DF X označíme F, pak:

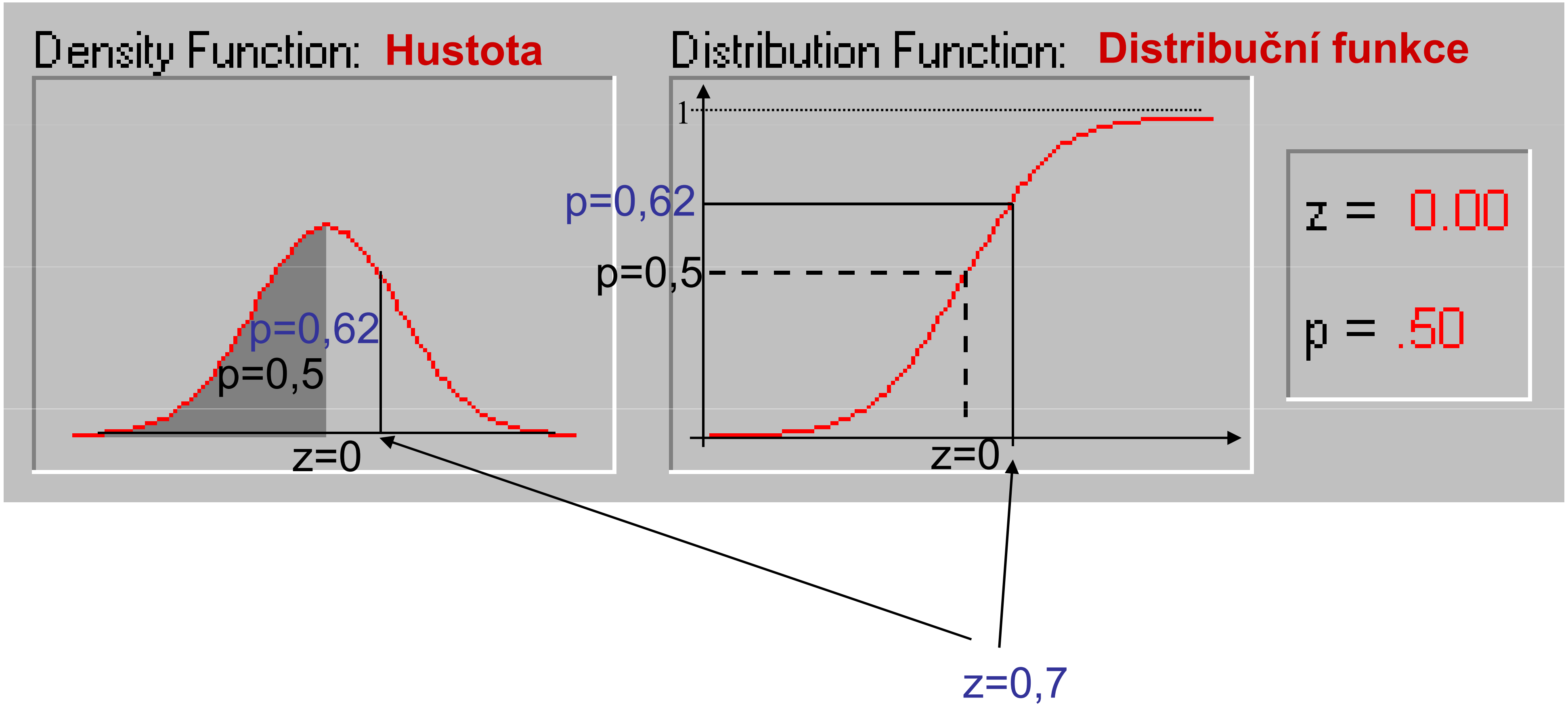
$$F(x) = P(X \leq x) \quad \text{pro každé } x \in \mathbf{R}$$

- ***funkce hustoty*** $f(x)$ NV X - každému intervalu čísel $[a,b]$, kde $a < b$ přiřazuje pravděpodobnost, že náhodná veličina X nabude hodnoty z tohoto intervalu, platí:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Vztah mezi:

Funkce hustoty a Distribuční funkce



Funkce hustoty versus funkce pravděpodobnosti 1

Hodnota FH není pravděpodobnost!!!

Pravděpodobnostní funkce $p(x)$ - každému x přiřazuje odpovídající pravděpodobnost: $p(x) = P(X = x)$
pro každé $x \in \boldsymbol{D}$ (diskrétní množina)

PF $p(x)$ splňuje vztahy:

$$\sum_{x \in D} p(x) = 1 \quad P(a \leq X \leq b) = \sum_{\substack{x=a \\ x \in D}}^{b} p(x)$$

Funkce hustoty versus funkce pravděpodobnosti 2

Funkce hustoty $f(x)$ - každému ***intervalu*** $[a,b]$ přiřazuje odpovídající pravděpodobnost, FH splňuje vztahy:

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \quad P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$$

Charakteristiky spojité náhodné veličiny

- **Charakteristiky polohy:**

  modus $Mod(X)$

  medián $Med(X)$

  střední hodnota $E(X)=\int_{-\infty}^{+\infty} xf(x)dx$

- **Charakteristiky variability:**

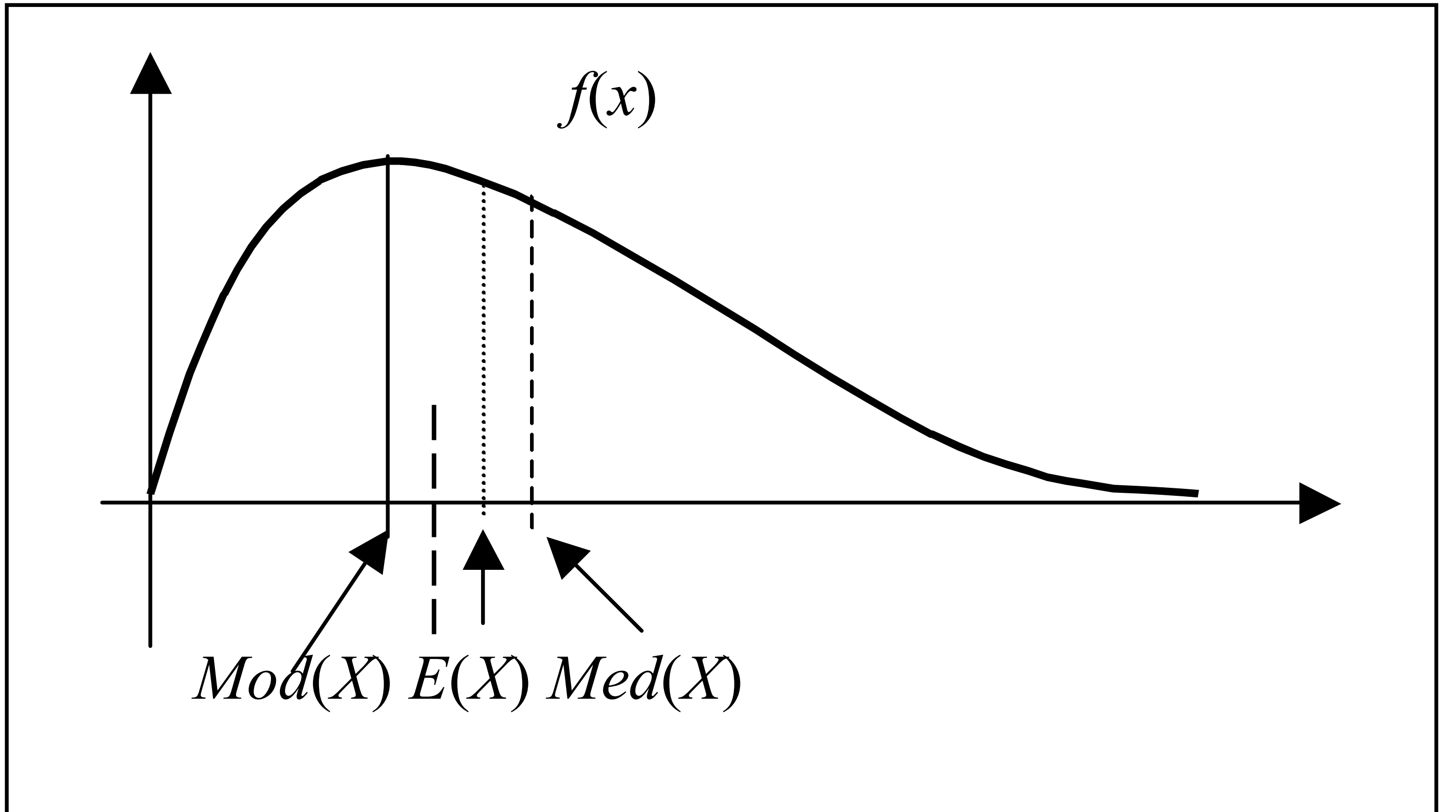
  rozptyl $Var(X)=\int_{-\infty}^{+\infty} [x-E(X)]^2 f(x)dx$

  směrodatná odchylka $\sigma(X)=\sqrt{Var(X)}$

Charakteristiky spojité náhodné veličiny

- Charakteristiky polohy:

  modus, medián, střední hodnota



Stejnoměrné rozdělení 1

Spojitá náhodná veličina X má ***stejnoměrné rozdělení*:**

- nabývá hodnot z intervalu $[a,b]$
- se stejnou pravděpodobností

***Funkce hustoty*:** $f(x) = \dfrac{1}{b-a}$ pro $x \in [a,b]$, jinak $f(x) = 0$

Pravděpodobnost: $c,d \in [a,b]$, $P(c \le X \le d) = \int_c^d f(x)dx = \dfrac{d-c}{b-a}$

Střední hodnota: $E(X) = \dfrac{a+b}{2}$

Rozptyl: $Var(X) = \dfrac{(b-a)^2}{12}, \quad \sigma(X) = \dfrac{(b-a)}{\sqrt{12}}$

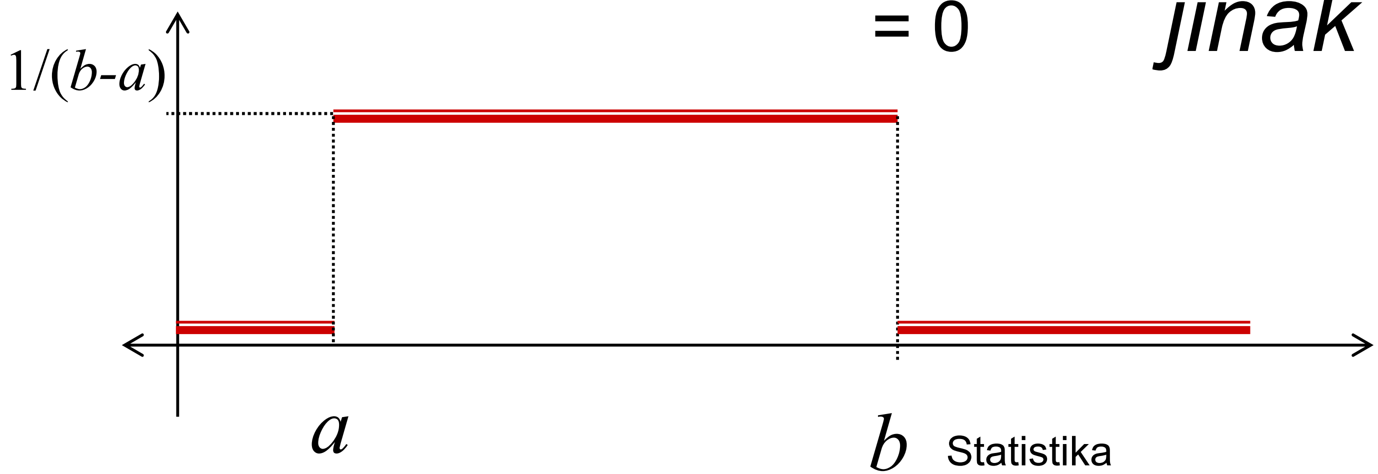
Stejnoměrné rozdělení 2

Spojitá náhodná veličina X má ***stejnoměrné rozdělení:***

- nabývá hodnot z intervalu $[a,b]$
- se stejnou pravděpodobností

s funkcí hustoty $f(x)=\dfrac{1}{b-a}$ *pro* x *z* $[a,b]$,

$= 0$ *jinak*



Příklad 1. „ NÁHČÍSLO“ v Excelu

- (Vložit-Funkce-Matematické: NÁHČÍSLO)
- Střední hodnota: $E(X) = (0+1)/2 = 0,5$
- Rozptyl: $Var(X) = (1 - 0)^2 /12 = 0,08$
- Sm. odchylka: $\sigma(X) = \sqrt{0,08} = 0,29$
- Pomocí lineární transformace (násobení a přičtení, resp. odečtení čísla) lze v Excelu generovat NV se stejnoměrným rozdělením v (libovolném) intervalu [*a* , *a* +*b*] vzorcem:

NV = *b**NÁHČÍSLO+ *a*

Příklad 2. Čekání na autobus

Autobusy odjíždějí z určité zastávky během dne pravidelně každých 15 minut. V náhodnou dobu přijdete na zastávku.

(a) Jaká je pravděpodobnost, že budete na autobus čekat dobu mezi 5 až 10 minutami?

(b) Jaká je pravděpodobnost, že budete čekat alespoň 12 minut?

(c) Stanovte střední hodnotu a směrodatnou odchylku doby čekání.

Příklad 2. Čekání na autobus - řešení

X je spojitá náhodná veličina s následující hustotou:

$$f(x) = \frac{1}{15} \quad \text{pro } 0 \le x \le 15$$

$$= 0 \qquad \text{jinde}$$

$$E(X) = \frac{0+15}{2} = 7{,}5$$

$$Var(X) = \frac{(15-0)^2}{12} = 18{,}75$$

Příklad 2. Čekání na autobus – pokračování

$$\boxed{P(c \le X \le d) = \frac{d-c}{b-a}}$$

(a) S využitím vzorce vypočítáme: $P(5<X<10)=(10\text{-}5)/(15\text{-}0) = 0{,}33$

(b) Analogicky obdržíme:

$P(12< X<15)= (15\text{-}12)/(15\text{-}0) = 0{,}2$

(c) $\sigma(X) = \sqrt{18{,}75} = 4{,}33$

Střední čekací doba je 7,5 minut, směrodatná odchylka je 4,33 minut

Normální rozdělení

Nejdůležitější rozdělení ve statistice!

Normální (Gaussovo) rozdělení pr-sti NV:

Způsobené kolísáním NV velkého počtu nepatrných a vzájemně nezávislých vlivů, které se skládají (sečítají)

Příklady:

(1) výsledky různých testů (body)

(2) výsledky měření rozměrů a hmotností (mm, cm, m, g, kg, t aj.)

Normální rozdělení 1

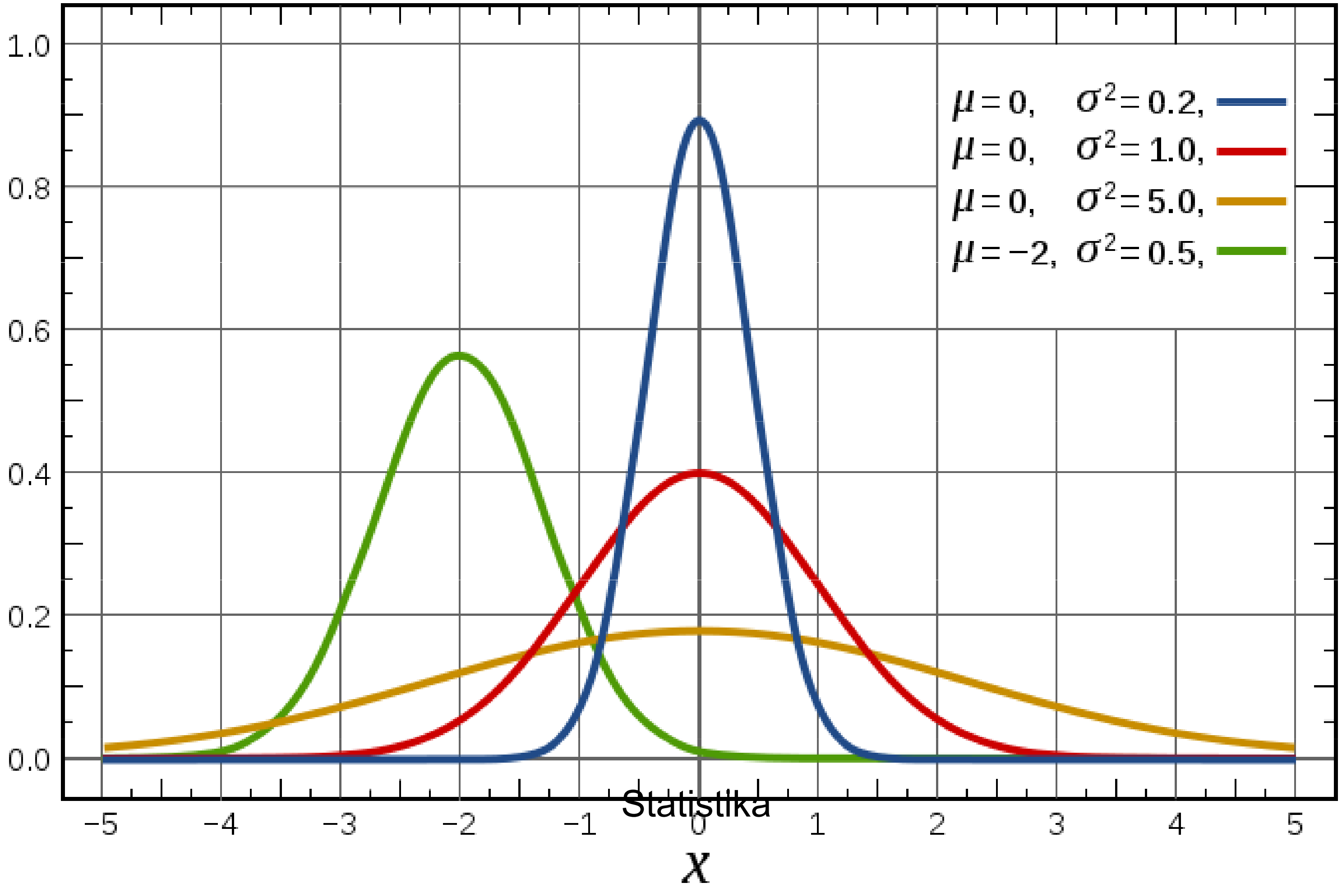
Funkce hustoty rozdělení pr-sti $f(x|\mu,\sigma^2)$:

$$\boxed{f(x\,|\,\mu,\sigma^2)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$$

$-\infty < x < +\infty,\ -\infty < \mu < +\infty,\ \sigma > 0$, kde μ a σ^2 se nazývají **parametry rozdělení**

Normální rozdělení 2

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Normální rozdělení 3

Charakteristiky:

Střední hodnota: $\boxed{E(X) = \mu}$

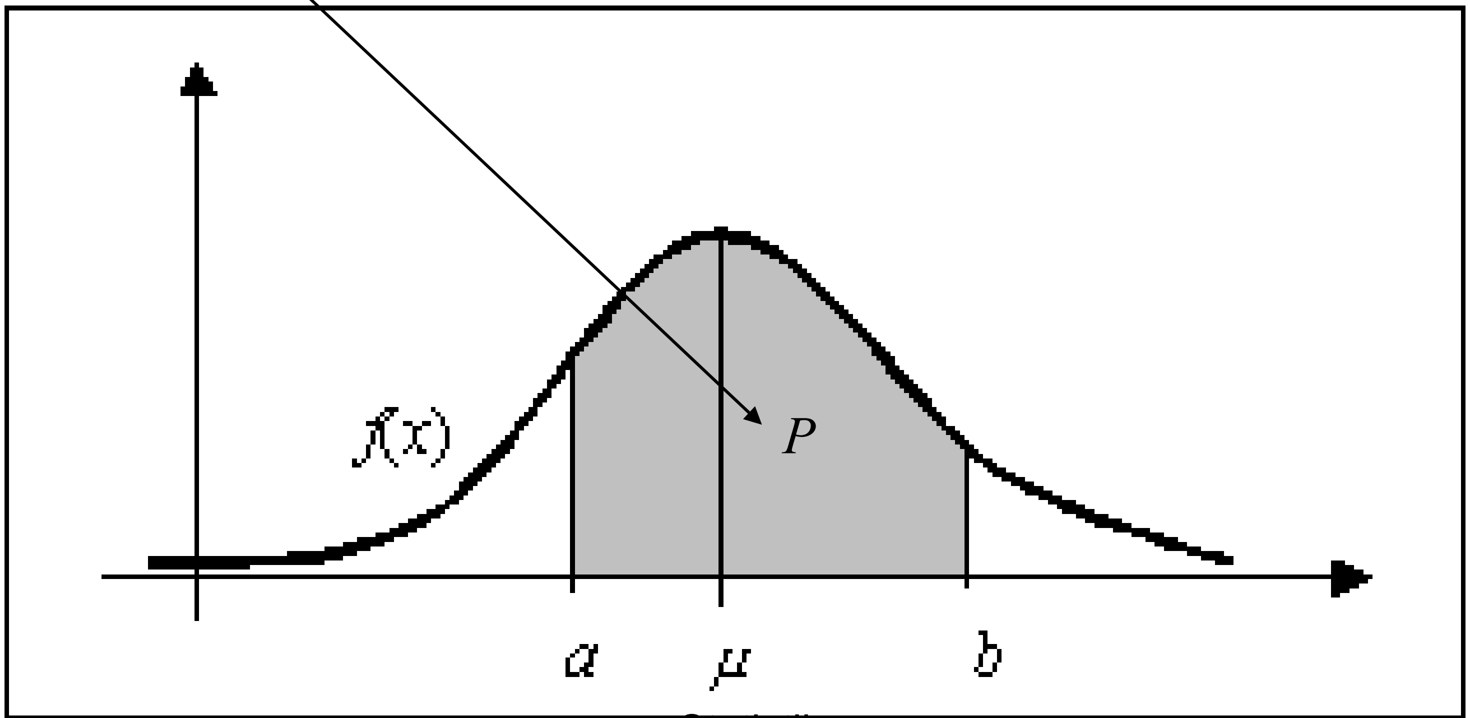
Rozptyl: $\boxed{Var(X) = \sigma^2}$

Směrodatná odchylka: $\boxed{\sigma(X) = \sigma}$

Normální rozdělení 4

Pravděpodobnost:

$$P(a \le X \le b) = \int_a^b f(x \mid \boldsymbol{\mu}, \boldsymbol{\sigma}^2)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$



Normální rozdělení 5

Normované normální rozdělení: μ = 0, σ^2 = 1

- Funkce hustoty: $f_N(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
- Pravděpodobnost (tabelována v tabulkách) viz skriptum:

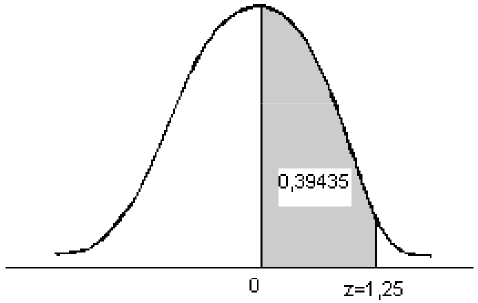
$$P(0 \le X \le z) = \int_0^z f_N(x)dx = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{x^2}{2}} dx$$

$$P(a \le X \le b) = P(0 \le X \le b) - P(0 \le X \le a)$$

- Pro $z \ge 3$ je pr-st téměř rovna 0,5 !

TABULKA 1

Plocha pod křivkou
normovaného normálního rozdělení $N(0,1)$



	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,00000	0,00399	0,00798	0,01197	0,01595	0,01994	0,02392	0,02790	0,03188	0,03586
0,1	0,03983	0,04380	0,04776	0,05172	0,05567	0,05962	0,06356	0,06749	0,07142	0,07535
0,2	0,07926	0,08317	0,08706	0,09095	0,09483	0,09871	0,10257	0,10642	0,10260	0,11409
0,3	0,11791	0,12172	0,12552	0,12930	0,13307	0,13683	0,14058	0,14431	0,14803	0,15173
0,4	0,15542	0,15910	0,16276	0,16640	0,17003	0,17364	0,18824	0,18082	0,18439	0,18793
0,5	0,19146	0,19497	0,19847	0,20194	0,20540	0,20884	0,21226	0,21566	0,21904	0,22240
0,6	0,22575	0,22907	0,23237	0,23565	0,23891	0,24215	0,24537	0,24857	0,25175	0,25490
0,7	0,25804	0,26115	0,26424	0,26730	0,27035	0,27337	0,27637	0,27935	0,28230	0,28524
0,8	0,28814	0,29103	0,29389	0,29673	0,29955	0,30234	0,30511	0,30785	0,31057	0,31327
0,9	0,31594	0,31859	0,32121	0,32381	0,32639	0,32894	0,33147	0,33398	0,36460	0,33891
1,0	0,34134	0,34375	0,34614	0,34850	0,35083	0,35314	0,35543	0,35769	0,35993	0,36214
1,1	0,36433	0,36650	0,36864	0,37076	0,37286	0,37493	0,37698	0,37900	0,38100	0,38298
1,2	0,38493	0,38686	0,38877	0,39065	0,39251	**0,39435**	0,39617	0,39796	0,39973	0,40147
1,3	0,40320	0,40490	0,40658	0,40824	0,40988	0,41149	0,41309	0,41466	0,41621	0,41774
1,4	0,41924	0,42073	0,42220	0,42364	0,42507	0,42647	0,42786	0,42922	0,43056	0,43189
1,5	0,43319	0,43448	0,43574	0,43699	0,43822	0,43943	0,44062	0,44179	0,44295	0,44408

Normální rozdělení 6

- Namísto NV X s normálním rozdělením s parametry μ, σ^2 uvažujeme transformovanou NV Z takto:

$$Z = \frac{x - \mu}{\sigma} \qquad (*)$$

- potom se funkce hustoty převede na hustotu normovaného normálního rozdělení transformaci (*) nazýváme ***normalizace***
- **V Excelu**: **NORMDIST**(x; Střed_hodn; Sm_odch; Součet)
  NORMINV(prst; střední; sm_odch)

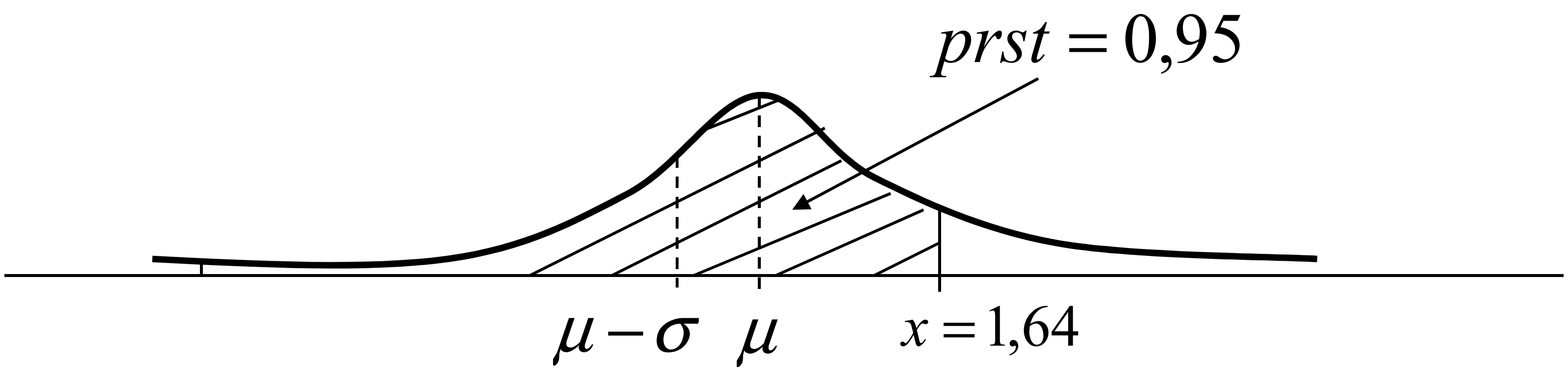
Normální rozdělení 7

v Excelu: **NORMDIST(x; Střed_hodn; Sm_odch; Součet)**

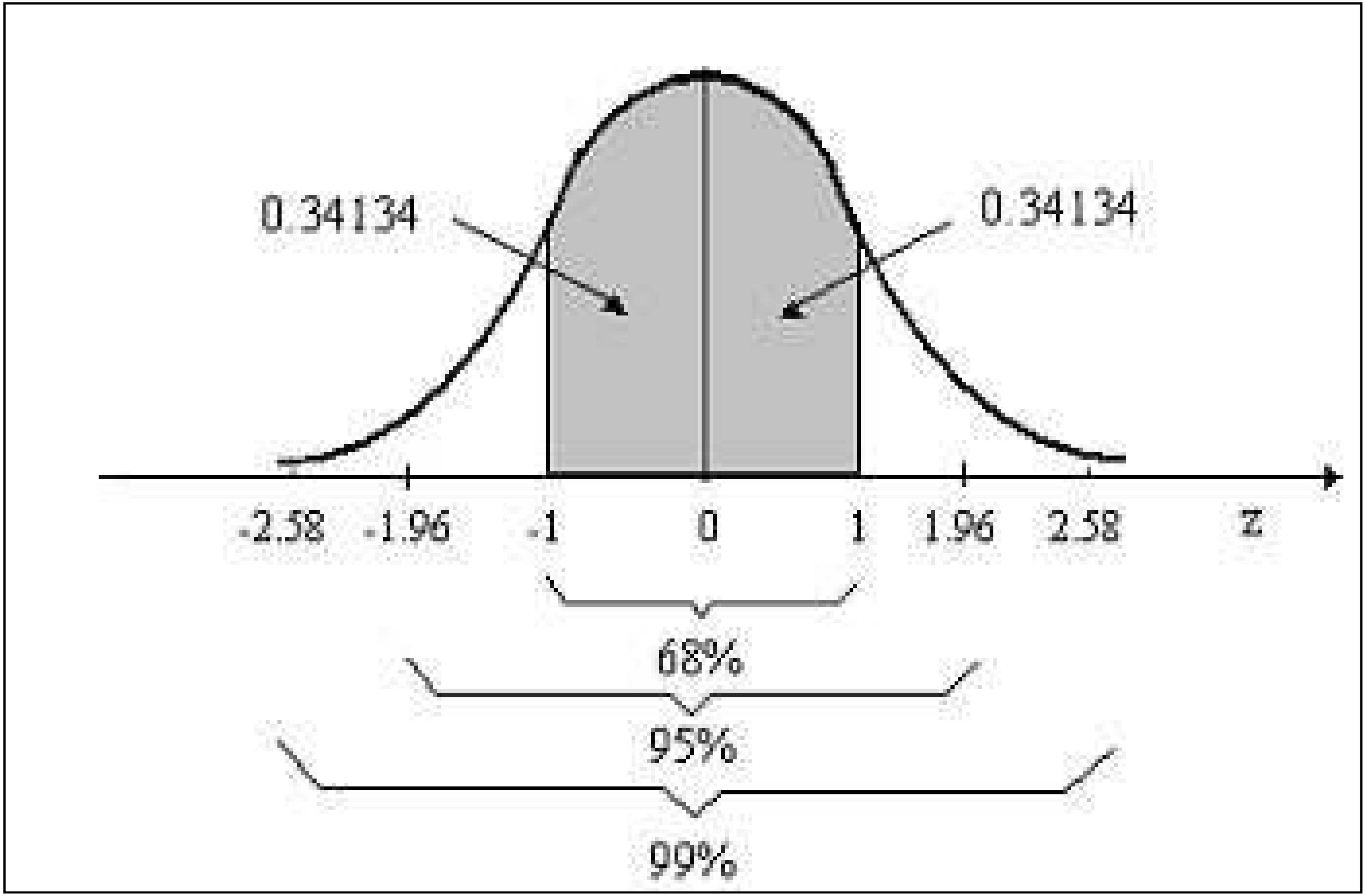
NORMINV(prst; střední; sm_odch)

NORMDIST(x; Střed_hodn; Sm_odch; Součet) = $P(X \leq \boldsymbol{x})$

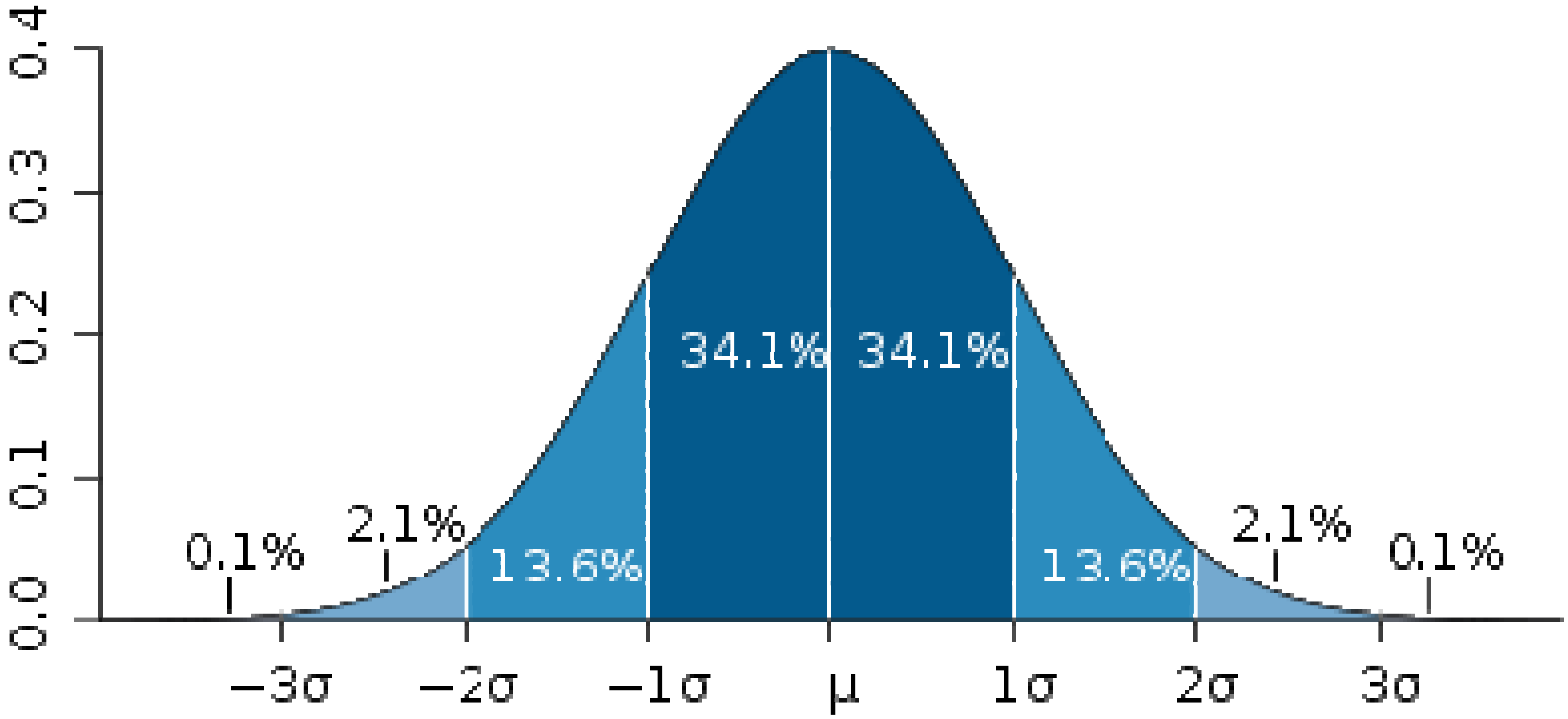
NORMINV(prst; střední; sm_odch) = $\boldsymbol{x}$ $P(X \leq \boldsymbol{x})$ = prst



Významné hodnoty normovaného normálního rozdělení $N(0,1)$



Významné hodnoty normovaného normálního rozdělení $N(\mu,\sigma^2)$



Příklad 3. Pomeranče

Jistý druh pomerančů má průměrnou hmotnost plodu μ =100 g se směrodatnou odchylkou σ =10 g.

(a) Jaká je pravděpodobnost, že náhodně vybraný plod bude mít hmotnost mezi 100g až 110g?

(b) Jaká je pravděpodobnost, že náhodně vybraný plod bude mít hmotnost větší než 120g?

Příklad 3. Pomeranče 2

Řešení: μ = 100, σ = 10

Ad **(a)** Normalizace: z_1=(100-100)/10 = 0

z_2=(110-100)/10 = 1

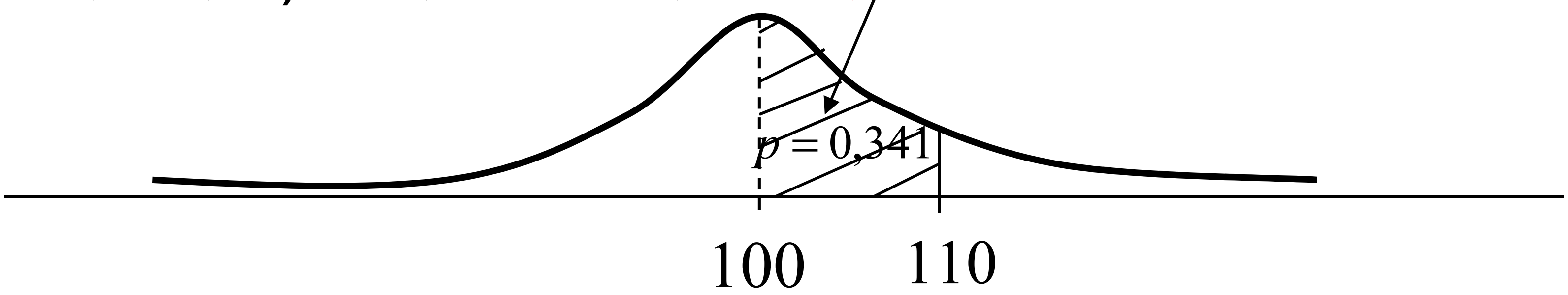
$$\boxed{Z = \frac{x-\mu}{\sigma}}$$

Platí: $P(100 \le X \le 110) = P(0 \le X \le 1)$

v tabulkách nalezneme: $P(0 \le X \le 1) = 0{,}341$

---

p = NORMDIST(110; 100; 10; 1) - NORMDIST(100; 100; 10; 1) = 0,841 – 0,5 = 0,341



Příklad 3. …Pomeranče 3

Řešení: μ = 100, σ= 10

Ad **(b)** Normalizace: $z = (120-100)/10 = 2$

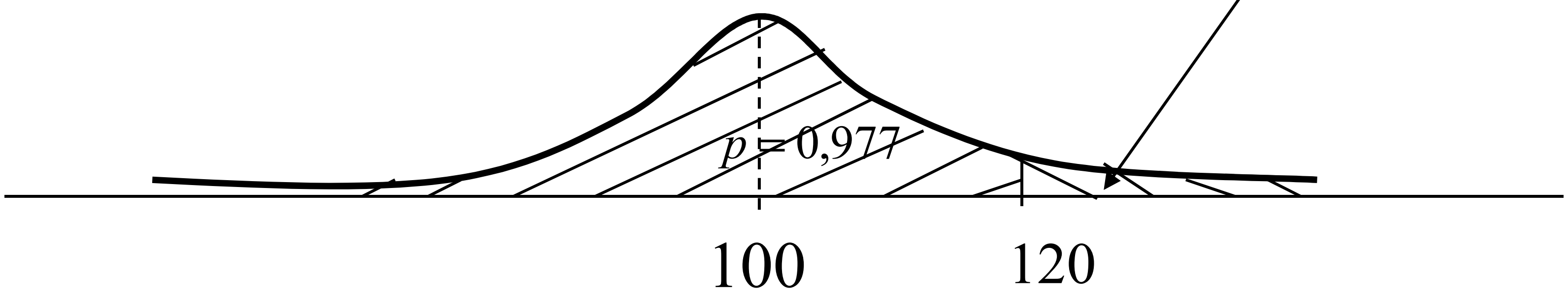
Podobně jako v (a):

$P(X \geq 120) = P(X \geq 2) = 0{,}5 - P(X < 2) =$

$0{,}5 - 0{,}477 = 0{,}023$

---

p = 1 - NORMDIST(120; 100; 10; 1) = 1 - 0,977 = 0,023



Exponenciální rozdělení 1

Exponenciální rozdělení pravděpodobnosti NV:

Exponenciální rozdělení slouží jako vhodný model pro výpočet pravděpodobnosti doby životnosti výrobků, čekacích dob v modelech hromadné obsluhy, apod.

Příklady: (1) doba pobytu ve frontě u přepážky

(2) doba obsluhy jednoho zákazníka

- Funkce hustoty rozdělení pravděpodobnosti $f(x|\ \delta)$:

$$\boxed{\begin{aligned} f(x) &= \frac{1}{\delta} e^{-\frac{x}{\delta}} \quad \text{pro } x > 0 \\ &= 0 \quad \text{jinak} \end{aligned}}$$

Přitom δ > 0 je parametr

Exponenciální rozdělení 2

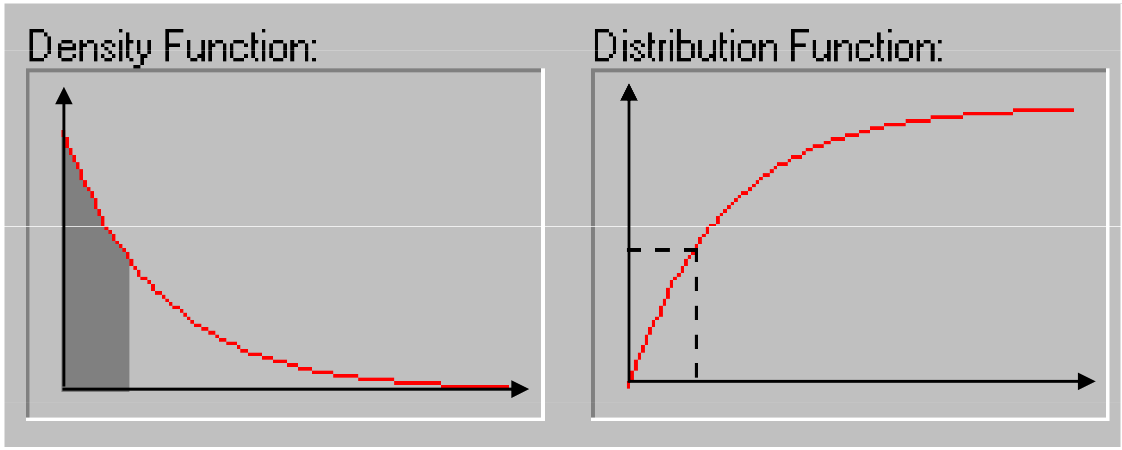
Charakteristiky:

Střední hodnota: $E(X) = \delta$

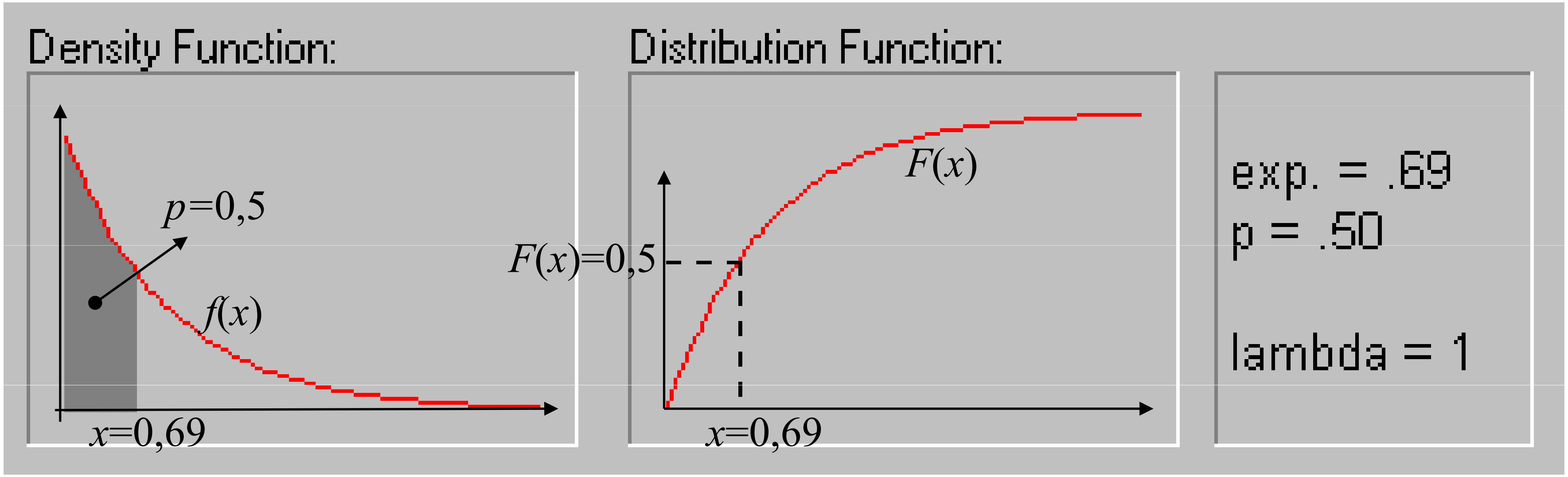
Rozptyl: $Var(X) = \delta^2$

Směrodatná odchylka: $\sigma(X) = \delta$ (= $\boldsymbol{E(X)}$!!!)

Pravděpodobnost: $P(a \leq X \leq b) = \int_a^b f(x)dx = e^{-\frac{a}{\delta}} - e^{-\frac{b}{\delta}}$



Exponenciální rozdělení 3



Funkce hustoty

Distribuční funkce

Exponenciální rozdělení: Příklad

Průměrná doba čekání u přepážky v bance je 5 min.

Jaká je pravděpodobnost, že zákazník bude čekat

(a) Právě 5 minut,

(b) Méně než 5 minut

(c) Více než 5 minut

(d) Více než 3 minuty a méně než 6 minut?

$$P(a \leq X \leq b) = \int_a^b f(x)dx = \left[-e^{-\frac{x}{\delta}} \right]_a^b = e^{-\frac{a}{\delta}} - e^{-\frac{b}{\delta}}$$

Příklad - řešení:

$$\boxed{P(a \le X \le b) = \left[-e^{-\frac{a}{\delta}}\right]_a^b = e^{-\frac{a}{\delta}} - e^{-\frac{b}{\delta}}}$$

Průměrná doba čekání u přepážky v bance je δ **= 5**.

(a) Právě 5 minut: $P(X = 5) = 0$!!! - spojité rozdělení,

(b) Více než 5 minut:

$$P(X \ge 5) = \left[-e^{-\frac{x}{5}}\right]_5^{+\infty} = e^{-\frac{5}{5}} - \lim_{x \to +\infty} e^{-\frac{x}{5}} - = e^{-1} - 0 \approx 0{,}368$$

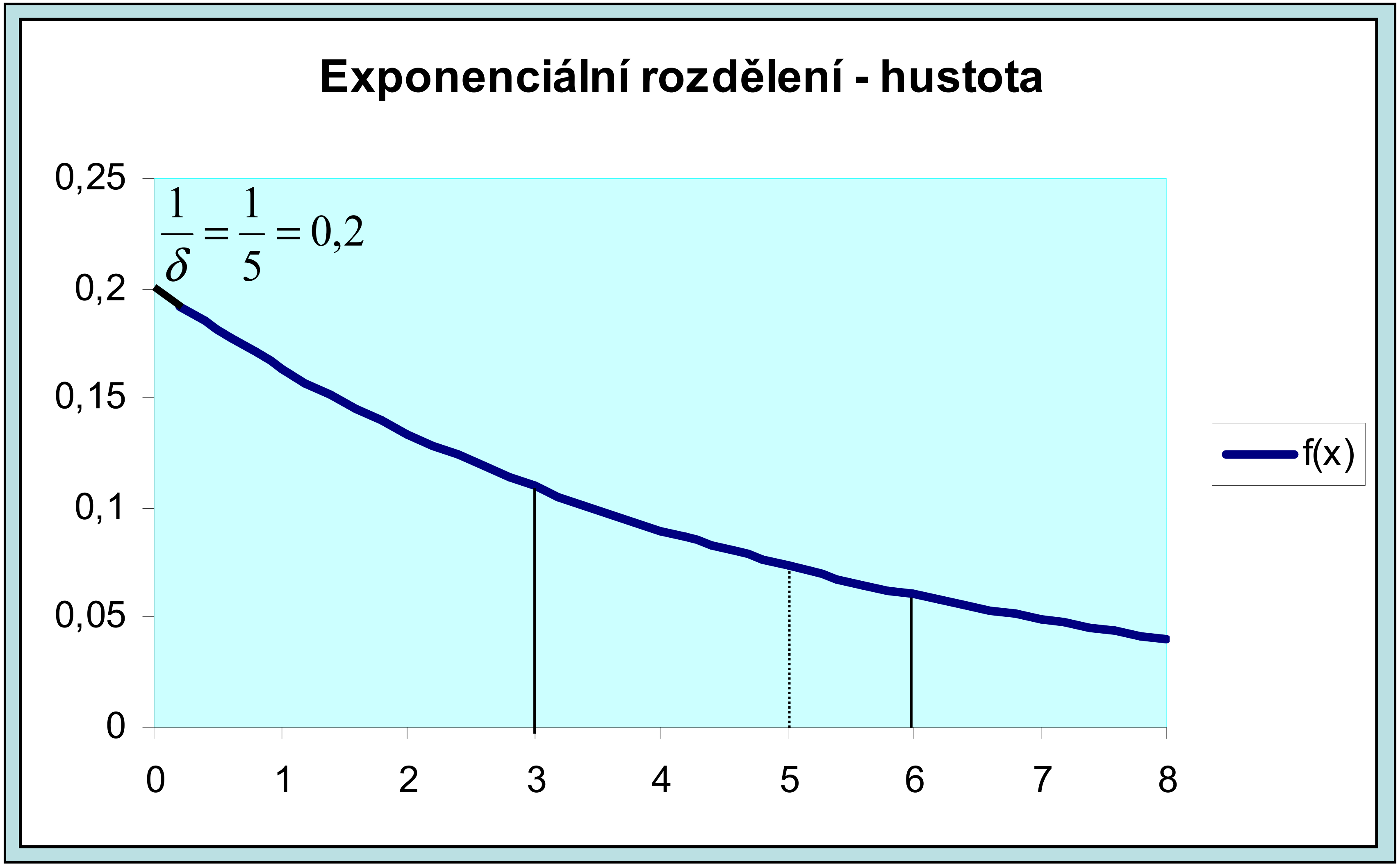
(c) Méně než 5 minut:

$$P(X \le 5) = \left[-e^{-\frac{x}{5}}\right]_0^{5} = e^{0} - e^{-1} \approx 1 - 0{,}368 = 0{,}632$$

(d) Více než 3 minuty a méně než 6 minut:

$$P(3 \le X \le 6) = \left[-e^{-\frac{x}{5}}\right]_3^{6} = e^{-\frac{3}{5}} - e^{-\frac{6}{5}} \approx 0{,}549 - 0{,}301 = 0{,}248$$

Exponenciální rozdělení - Příklad: graf



Shrnutí

- Co je to NV: číselné hodnoty + pravděpodnosti
- Spojitá NV: individuální čís. hod. + funkce hustoty pr-sti
- Charakteristiky NV (polohy - 3, variability - 2)
- NV se stejnoměrným rozdělením pr-sti
- NV s normálním rozdělením pr-sti
- NV s normovaným normálním rozdělením pr-sti
- NV s exponenciálním rozdělením