Základy pravděpodobnosti a matematické statistiky

5. Spojité pravděpodobnostní modely

FAKULTA STROJNÍ ČVUT V PRAZE

prof. RNDr. Gejza Dohnal, CSc.

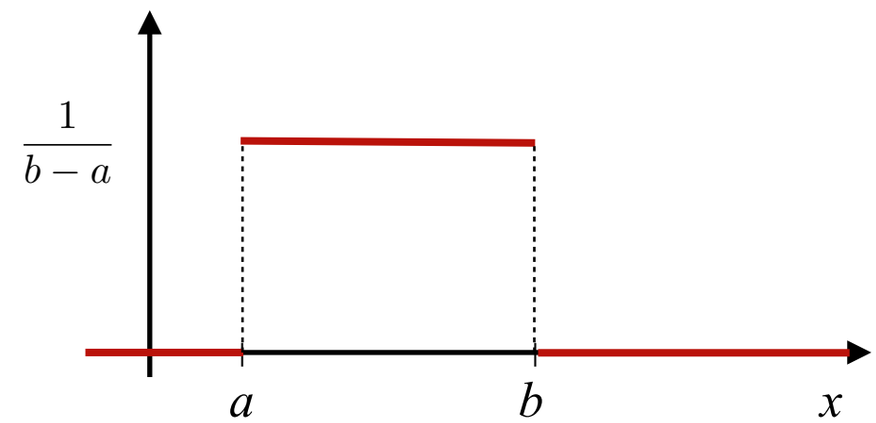
ak. rok 2022/2023

5. Spojité pravděpodobnostní modely

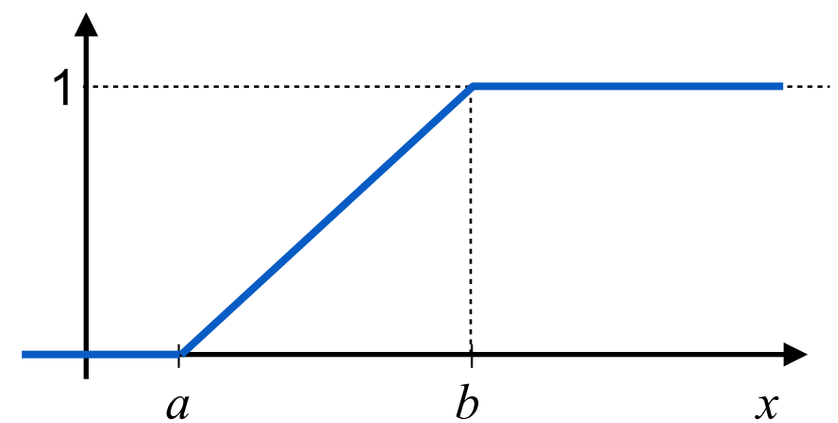
6) Model rovnoměrného rozdělení na intervalu

Model experimentu, při němž mohou nastat výsledky kdekoli v nějakém omezeném intervalu, přičemž nemáme důvod preferovat některou jeho část.

$$f(x) = \begin{cases} 0 & x < a, \\ \frac{1}{b-a} & x \in \langle a; b \rangle, \\ 0 & x > b. \end{cases}$$



$$F(x) = \begin{cases} 0 & x < a, \\ \frac{x-a}{b-a} & x \in \langle a; b \rangle, \\ 1 & x > b. \end{cases}$$



$$E(X) = \int_{-\infty}^{+\infty} x f(x) \mathrm{d}x = \int_a^b \frac{x}{b-a} \mathrm{d}x$$

$$= \frac{\left[x^2\right]_a^b}{2(b-a)} = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}$$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) \mathrm{d}x = \int_a^b \frac{x^2}{b-a} \mathrm{d}x = \frac{b^3 - a^3}{3(b-a)} = \frac{b^2 + ab + a^2}{3}$$

$$Var(X) = E(X^2) - (E(X))^2 = \frac{b^2 + ab + a^2}{3} - \frac{b^2 + 2ab + a^2}{4} = \frac{(b-a)^2}{12}$$

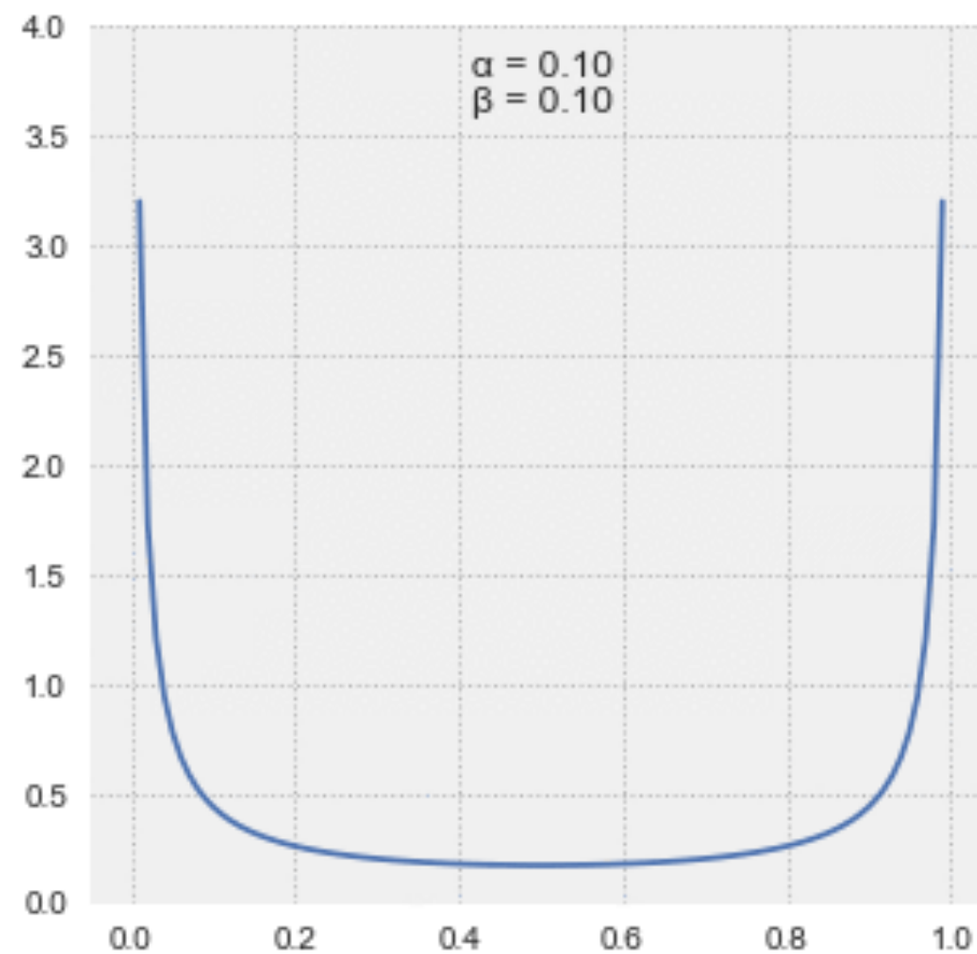
7) Model náhodného chování na omezeném intervalu

Model experimentu, při němž mohou nastat výsledky v nějakém omezeném intervalu, přičemž máme důvod preferovat některou jeho část.

- Zpravidla se uvádí pro interval ⟨0;1⟩ (pro jiné intervaly musíme nejprve provést transformaci.
- Často se používá pro modelování náhodného chování procent a podílů.

$$f(x) = \begin{cases} 0 & x < 0, \\ \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1 - x^{\beta-1}) & x \in \langle 0; 1 \rangle, \\ 0 & x > 1. \end{cases}$$

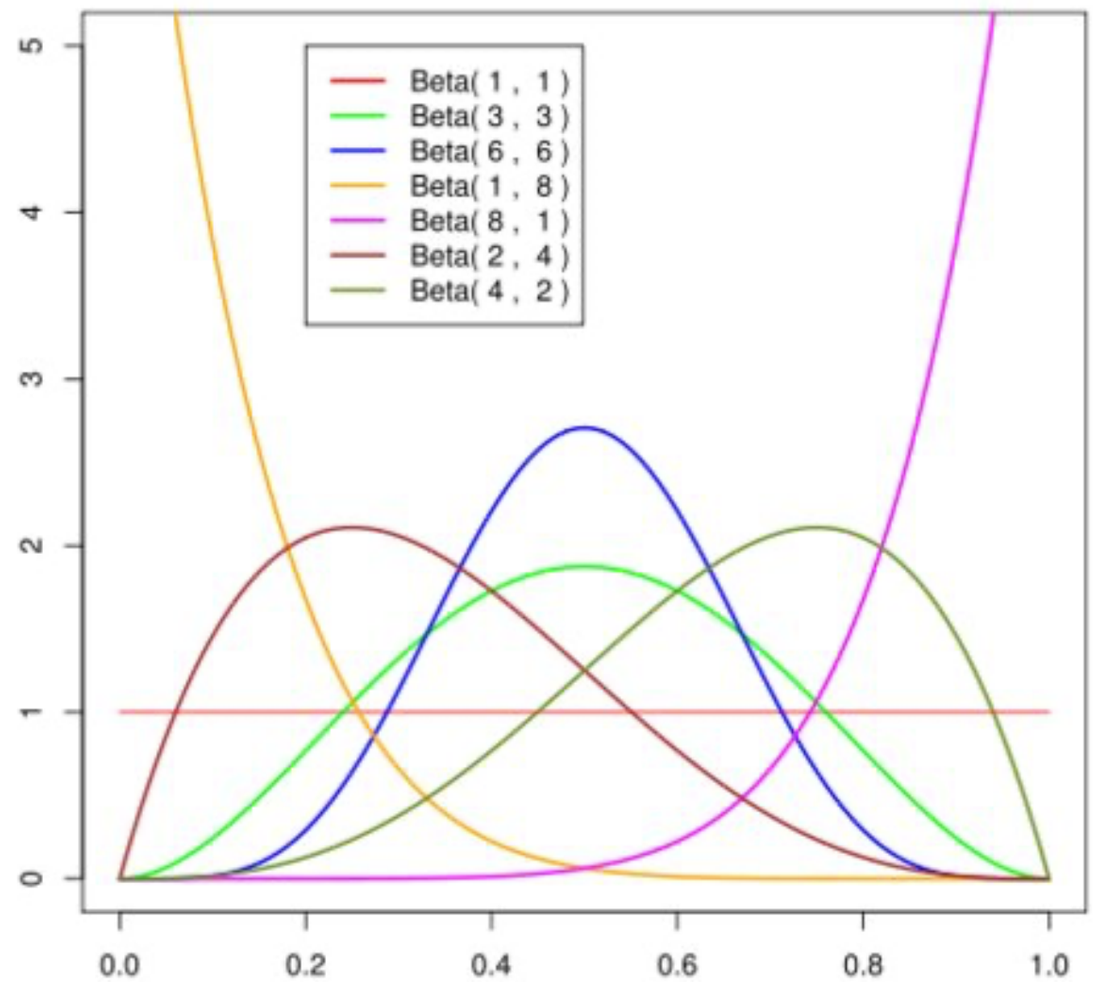
kde $B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} \mathrm{d}t$



7) Model náhodného chování na omezeném intervalu

Model experimentu, při němž mohou nastat výsledky v nějakém omezeném intervalu, přičemž máme důvod preferovat některou jeho část.

- Zpravidla se uvádí pro interval ⟨0;1⟩ (pro jiné intervaly musíme nejprve provést transformaci.
- Často se používá pro modelování náhodného chování procent a podílů.



$$f(x) = \begin{cases} 0 & x < 0, \\ \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1 - x^{\beta-1}) & x \in \langle 0; 1 \rangle, \\ 0 & x > 1. \end{cases}$$

kde $B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} \mathrm{d}t$

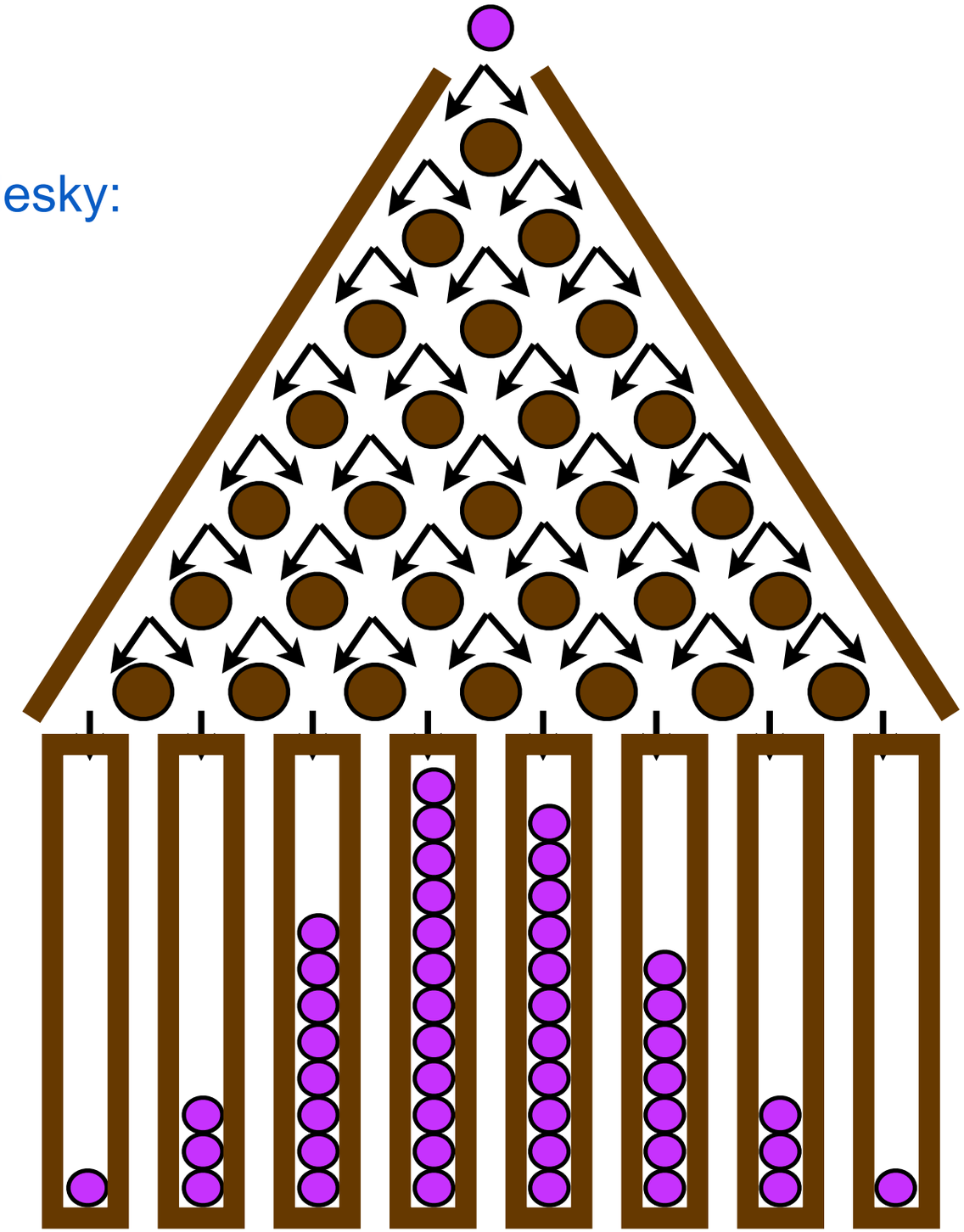
$$E(X) = \frac{\alpha}{\alpha + \beta}$$

$$Var(X) = \frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2}$$

7) Model normálního rozdělení

- 1733, Abraham de Moivre: při pokusech s mincemi se pokoušel nahradit (binomické) rozdělení spojitou křivkou

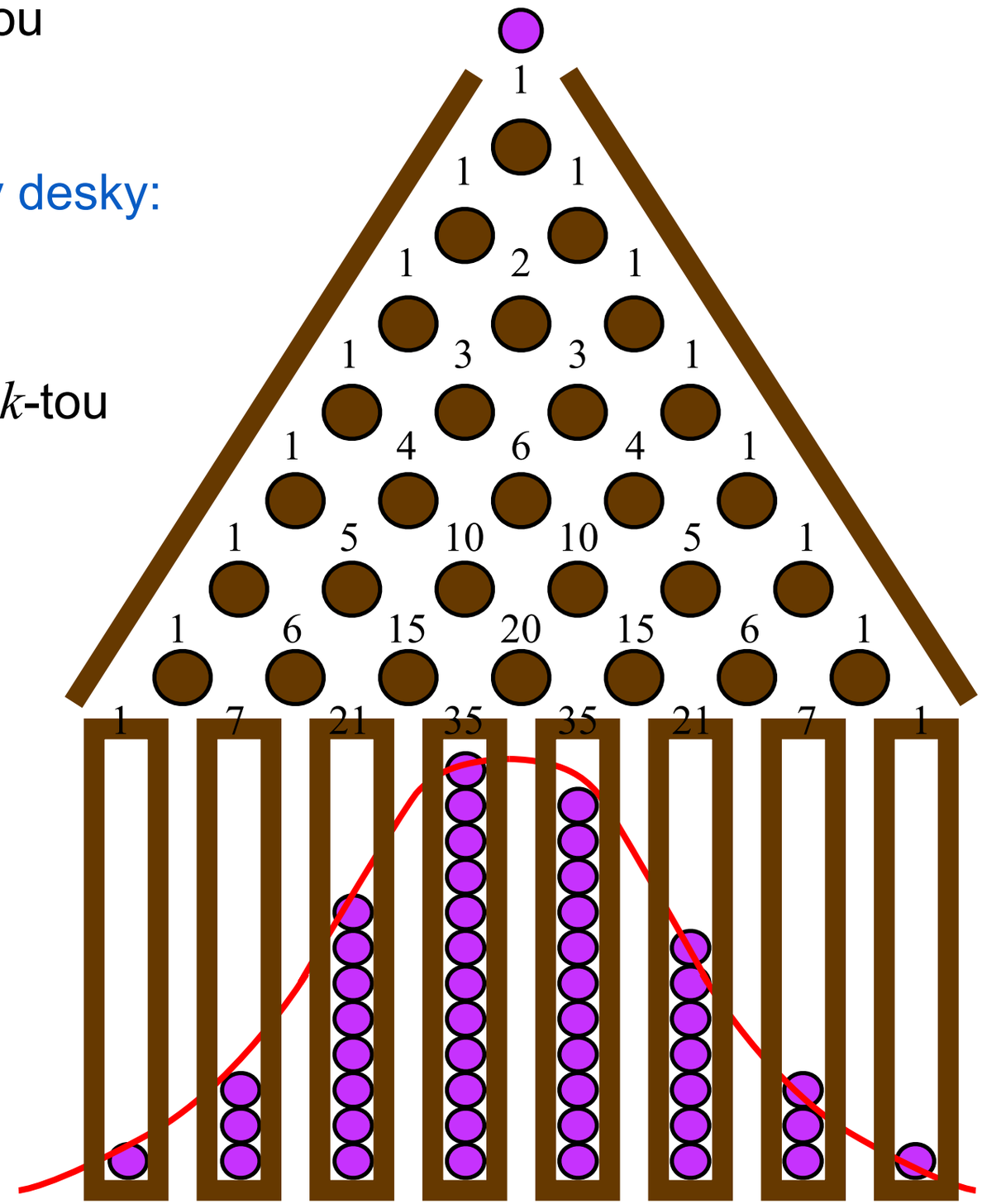
model Galtonovy desky:



7) Model normálního rozdělení

- 1733, Abraham de Moivre: při pokusech s mincemi se pokoušel nahradit (binomické) rozdělení spojitou křivkou

model Galtonovy desky:



Pascalův trojúhelník: pro N-tý řádek a k-tou pozici platí:

$$n_k = \binom{N}{k} = \frac{N!}{k!(N-k)!}$$

Moivre navrhl rovnici: $y = e^{-\frac{x^2}{2}}$

$$\boxed{y = K e^{-\frac{(x-a)^2}{2b^2}}}$$

7) Model normálního rozdělení

- 1733, Abraham de Moivre: při pokusech s mincemi se pokoušel nahradit (binomické) rozdělení spojitou křivkou

- přelom 18. a 19. století: rozvoj astronomie => rozdíly v astronomických měřeních - která hodnota je správná?

  Úlohu řešili Karl Friedrich Gauss a Piere Laplace a charakterizovali četnosti odchylek naměřených hodnot od vypočtených tzv. “křivkou chyb měření”

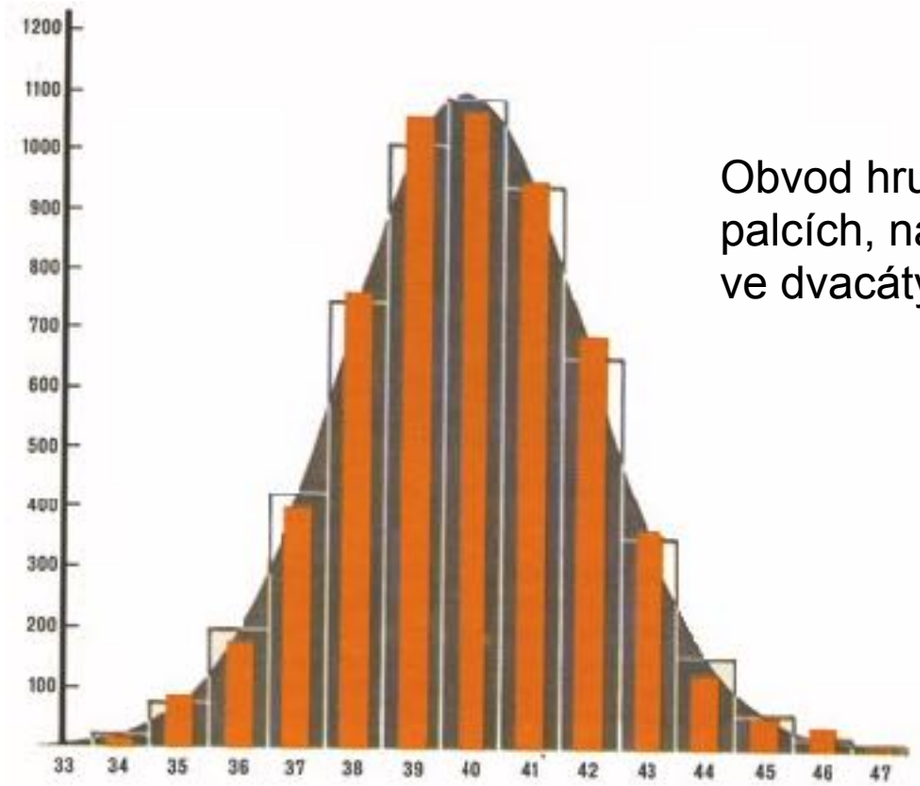
  → Gaussova křivka

- dvacátá léta 19. století: potřeba obléci anglickou armádu – bylo třeba nalézt rozměry „normálního“ vojáka. Belgický matematik Adolphe Qételet prováděl antropometrická měření (mimo jiné zavedl i tzv. BMI)

  → Normální rozdělení

Pro naše aplikace má tento model nejblíže k modelu odchylek (chyb) měření:

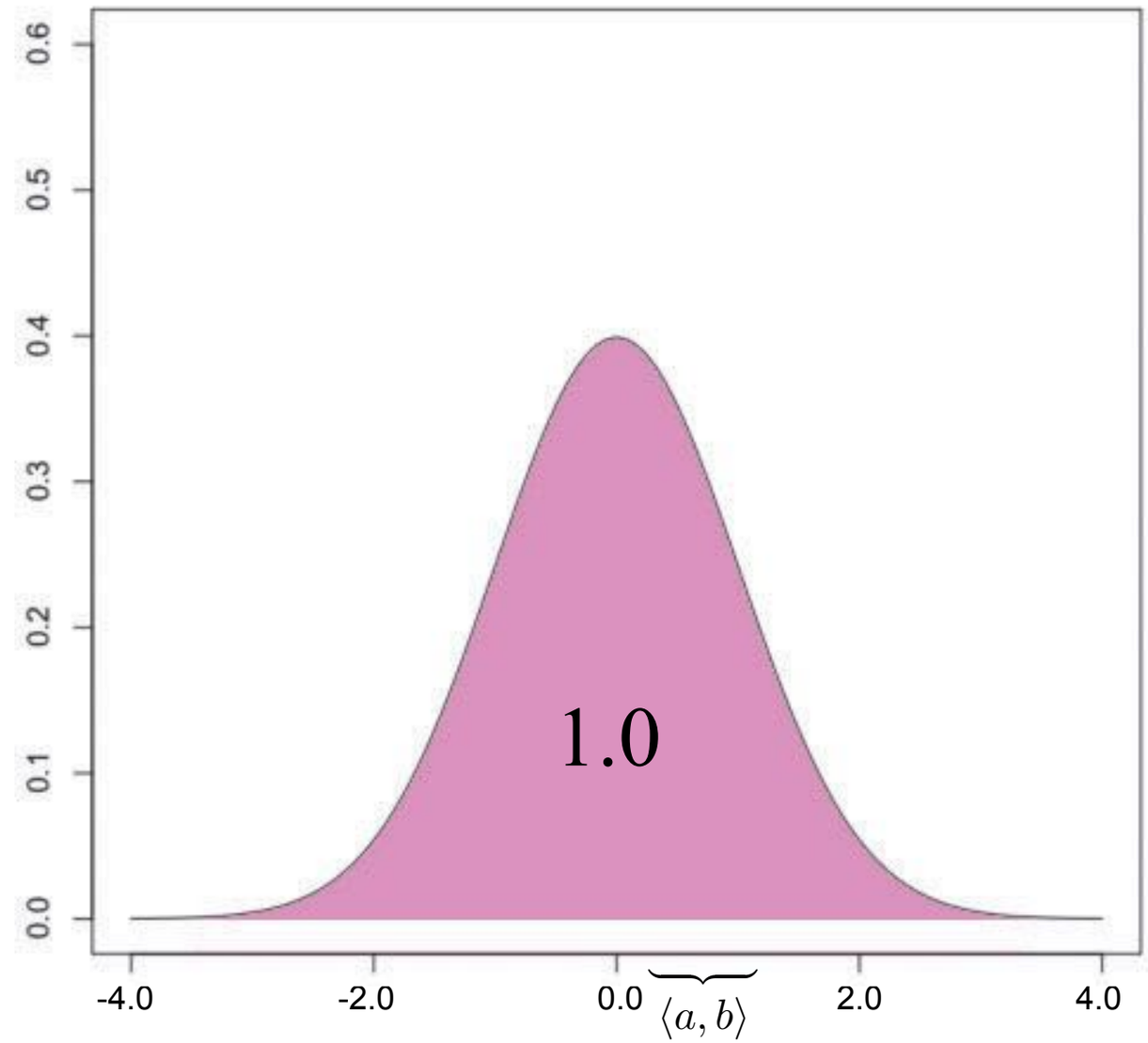
- symetrické kolem nuly
- „malé“ jsou časté
- „velké“ spíš vyjímečně



Obvod hrudi skotských vojáků v palcích, naměřený Quételetem ve dvacátých letech 19. století

7) Model normálního rozdělení

- základní (standardní) tvar hustoty: $$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in R$$



Standardní normální rozdělení

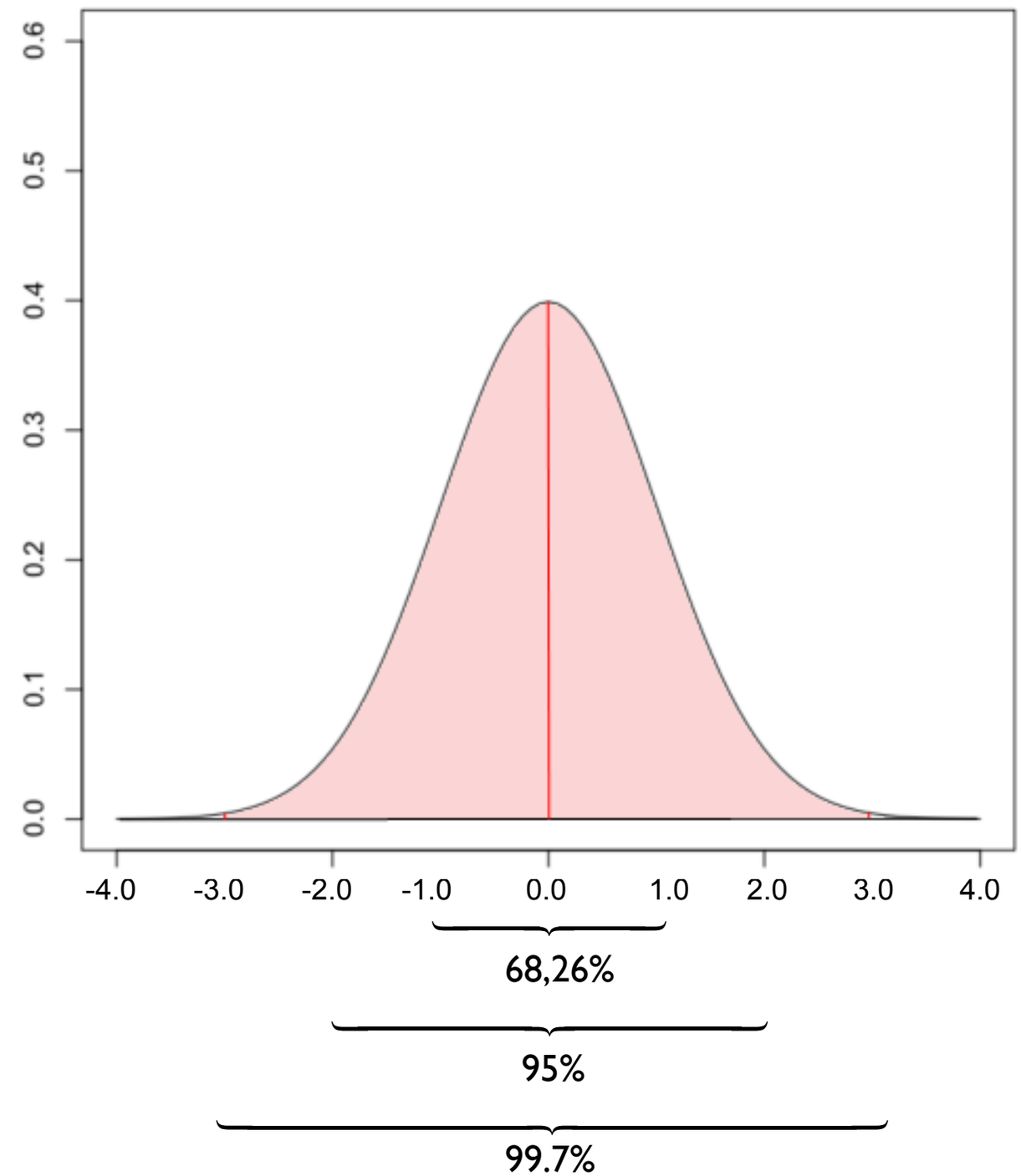
$$X \sim N(0,1) \qquad X \in (-\infty, +\infty)$$

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

$$\boxed{P(X \leq x) = \int_{-\infty}^{x} f(u)\mathrm{d}u = \Phi(x)}$$

7) Model normálního rozdělení

- základní (standardní) tvar hustoty: $$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in R$$



Standardní normální rozdělení

$$X \sim N(0,1) \qquad X \in (-\infty, +\infty)$$

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

$$P(X \leq x) = \int_{-\infty}^{x} f(u)\mathrm{d}u = \Phi(x)$$

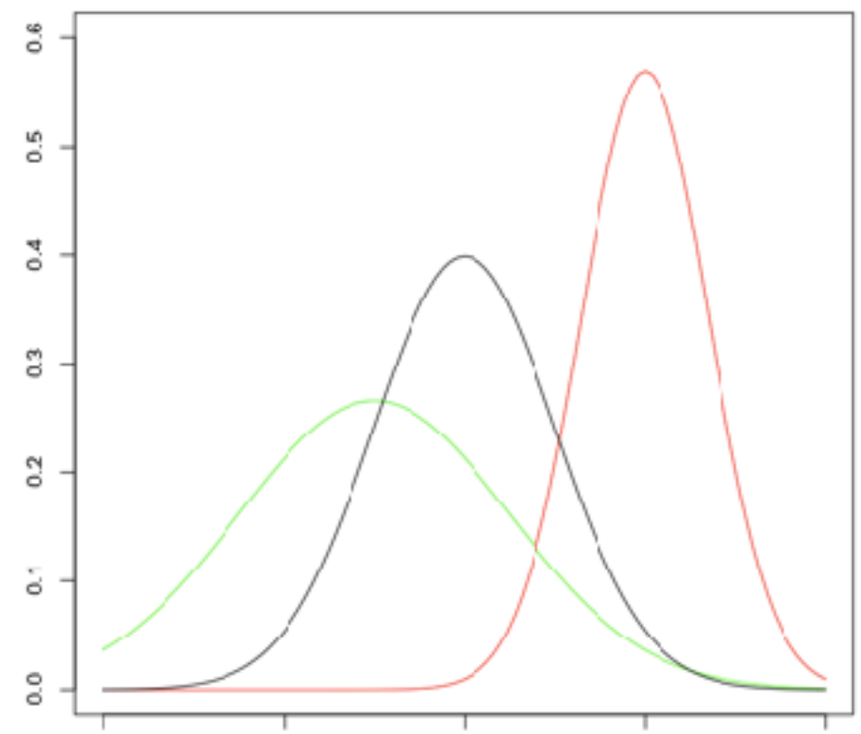
$$P(-1 \leq X \leq 1) = 0.6826$$

$$P(-2 \leq X \leq 2) = 0.9545$$

$$P(-3 \leq X \leq 3) = 0.9973$$

7) Model normálního rozdělení

- základní (standardní) tvar hustoty: $$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in R$$
- doplníme parametr měřítka $\sigma > 0$: $$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$
- doplníme parametr posunutí (polohy) $\mu \in R$: $$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$$



Obecné normální (Gaussovo) rozdělení $N(\mu,\sigma^2)$

7) Model normálního rozdělení

- základní (standardní) tvar hustoty N(0,1):

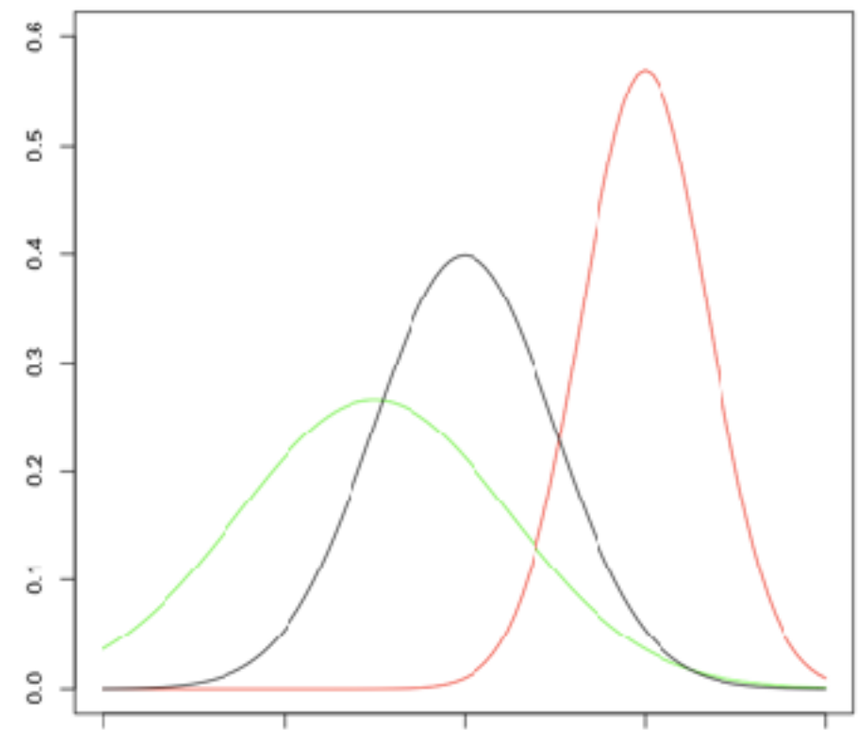
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in R$$

- doplníme parametr měřítka $\sigma > 0$:

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

- doplníme parametr posunutí (polohy) $\mu \in R$:

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



Obecné normální (Gaussovo) rozdělení N(μ,σ²)

$$X \sim N(0,1) \;\Rightarrow\; Y = \sigma X + \mu \sim N(\mu, \sigma^2)$$

a naopak:

$$Y \sim N(\mu, \sigma^2) \Rightarrow X = \frac{Y-\mu}{\sigma} \sim N(0,1)$$

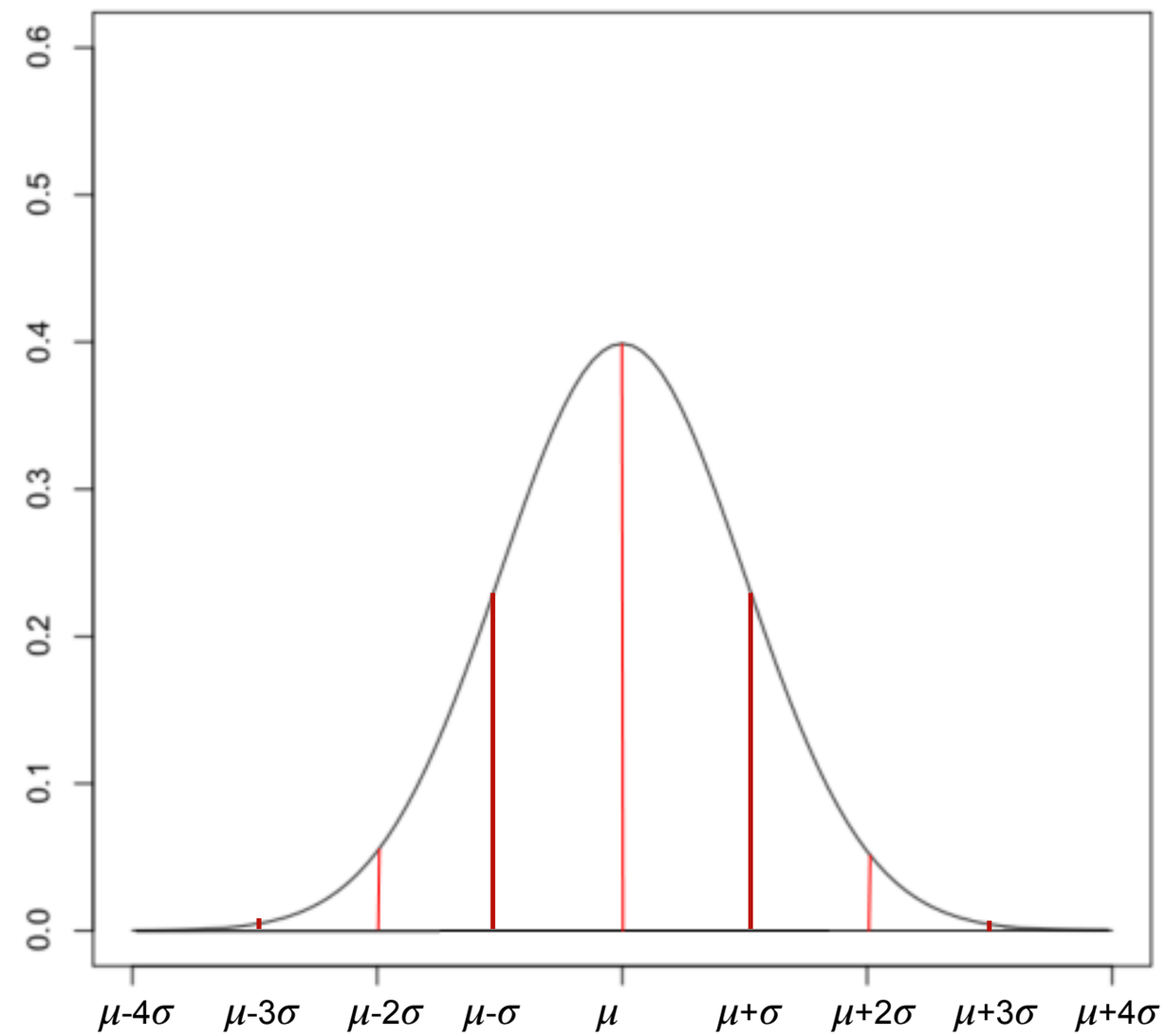
- tzv. standardizace

$$P(a \le Y \le b) = P(a-\mu \le Y-\mu \le b-\mu) = P\left(\frac{a-\mu}{\sigma} \le \frac{Y-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right)$$

$$= \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} f(u)\mathrm{d}u = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \quad \Rightarrow \quad \boxed{F(y) = \Phi\left(\frac{y-\mu}{\sigma}\right)}$$

7) Model normálního rozdělení

- obecný tvar hustoty: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$



$$P(-1 \leq X \leq 1) = 0.6826$$

$$P(-2 \leq X \leq 2) = 0.9545$$

$$P(-3 \leq X \leq 3) = 0.9973$$

$$P(-1 \leq X \leq 1) = P(-\sigma \leq \sigma X \leq \sigma) = P(-\sigma + \mu \leq \sigma X + \mu \leq \sigma + \mu)$$
$$= P(\mu - \sigma \leq Y \leq \mu + \sigma) = 0.6826$$

$$P(-2 \leq X \leq 2) = P(-2\sigma \leq \sigma X \leq 2\sigma) = P(-2\sigma + \mu \leq \sigma X + \mu \leq 2\sigma + \mu)$$
$$= P(\mu - 2\sigma \leq Y \leq \mu + 2\sigma) = 0.9545$$

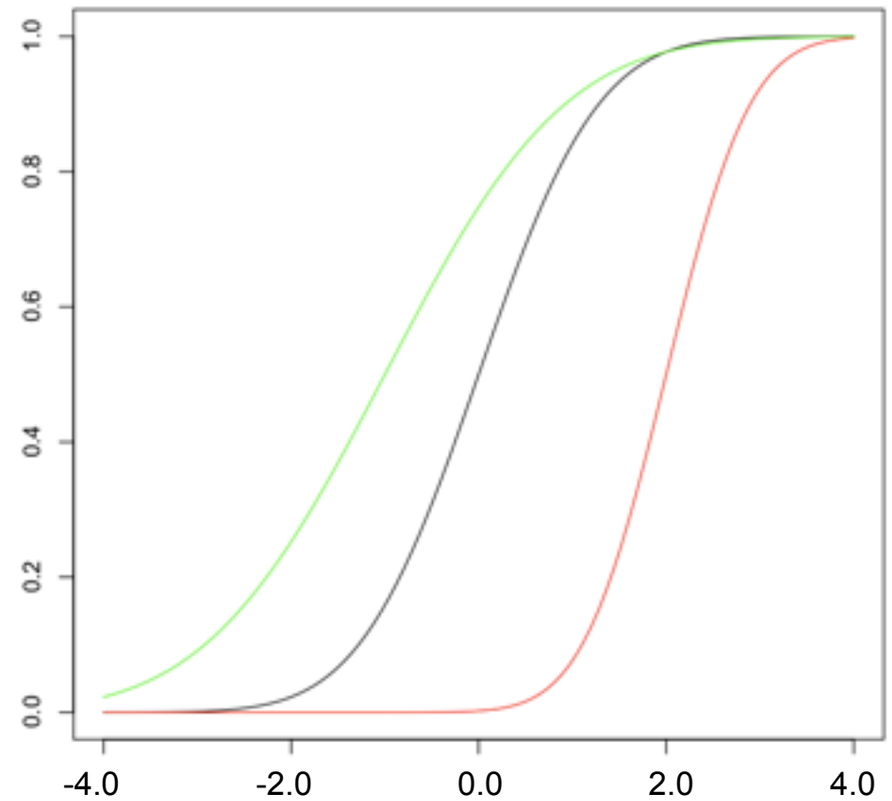
$$P(-3 \leq X \leq 3) = P(-3\sigma \leq \sigma X \leq 3\sigma) = P(-3\sigma + \mu \leq \sigma X + \mu \leq 3\sigma + \mu)$$
$$= P(\mu - 3\sigma \leq Y \leq \mu + 3\sigma) = 0.9973$$

7) Model normálního rozdělení

- obecný tvar hustoty: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

- distribuční funkce

$$P(X \leq x) = \int_{-\infty}^{x} f(u)\mathrm{d}u$$



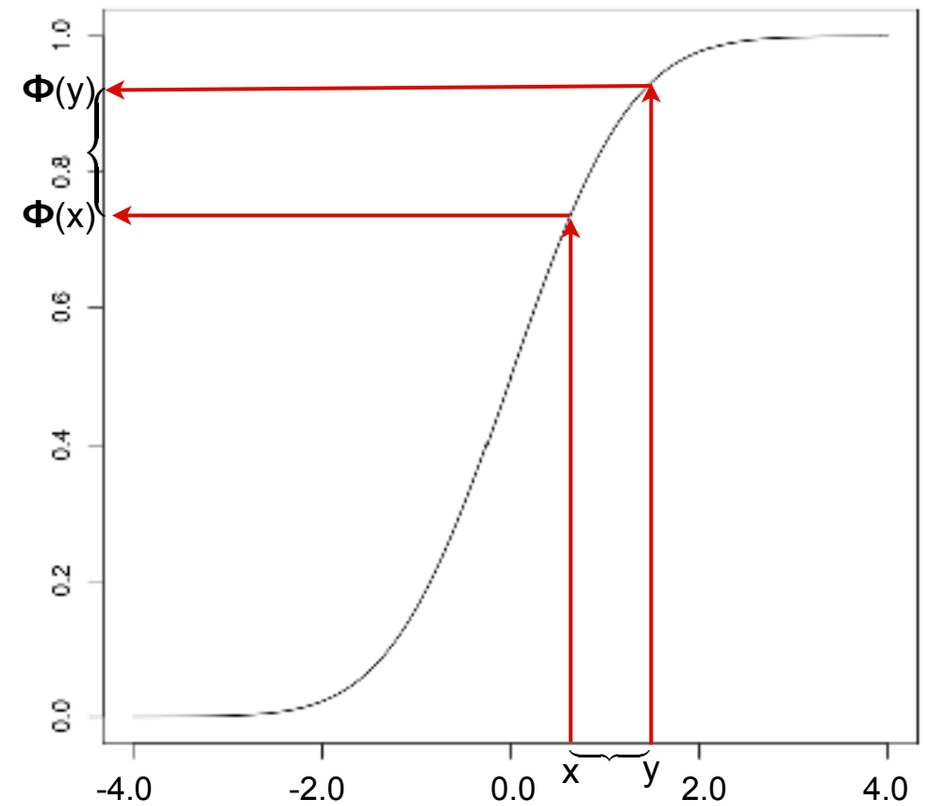
7) Model normálního rozdělení

- základní (standardní) tvar hustoty N(0,1):

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

- distribuční funkce

$$P(X \leq x) = \int_{-\infty}^{x} f(u)\mathrm{d}u = \Phi(x)$$



Jakou hodnotu náhodná veličina X nepřekročí s pravděpodobností p?

S jakou pravděpodobností bude budoucí hodnota náhodné veličiny X v intervalu $\langle x,y \rangle$?

$$P(a \leq X \leq b) = \int_a^b f(x)dx = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx = \Phi(y) - \Phi(x)$$

Jakých hodnot bude náhodná veličina X nabývat s předem danou pravděpodobností p?

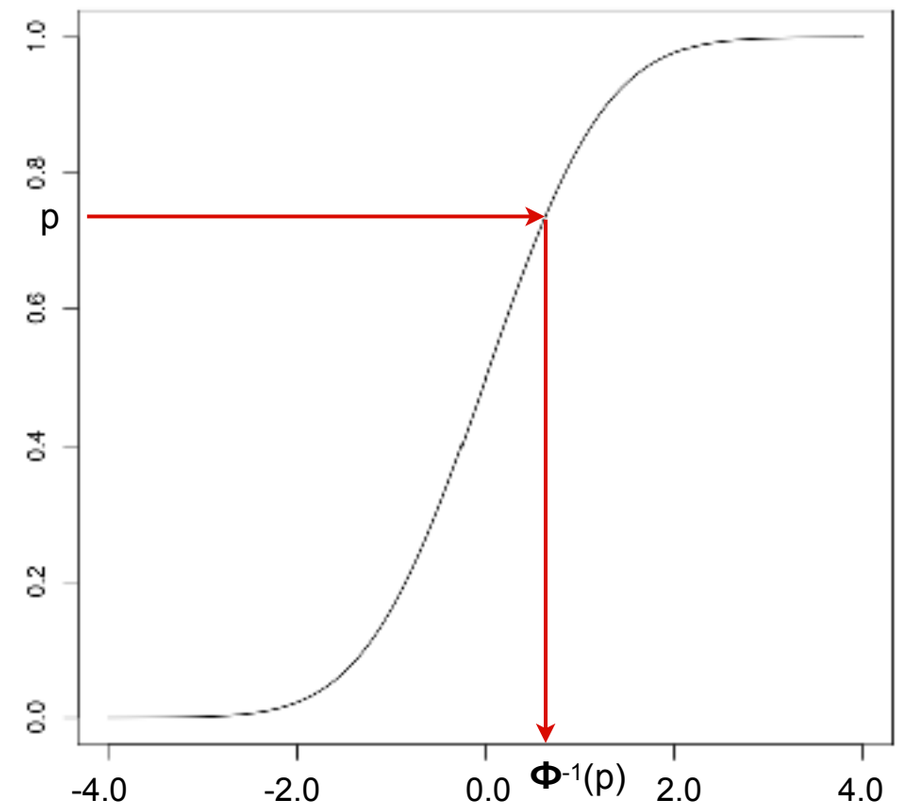
7) Model normálního rozdělení

- základní (standardní) tvar hustoty N(0,1):

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

- distribuční funkce

$$P(X \leq x) = \int_{-\infty}^{x} f(u)\mathrm{d}u = \Phi(x)$$



Jakou hodnotu náhodná veličina X nepřekročí s pravděpodobností p?

S jakou pravděpodobností bude budoucí hodnota náhodné veličiny X v intervalu $\langle x,y \rangle$?

$$P(a \leq X \leq b) = \int_a^b f(x)dx = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx = \Phi(y) - \Phi(x)$$

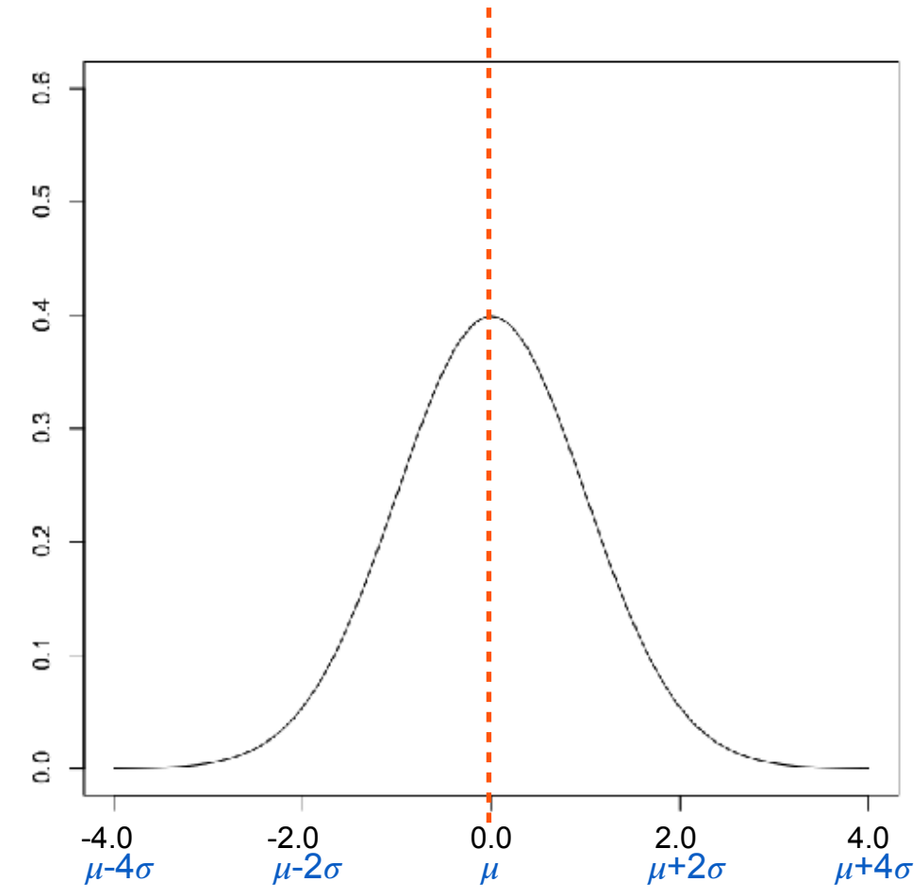
Jakých hodnot bude náhodná veličina X nabývat s předem danou pravděpodobností p?

- kvantilová funkce: $x_p = \Phi^{-1}(p)$

7) Model normálního rozdělení

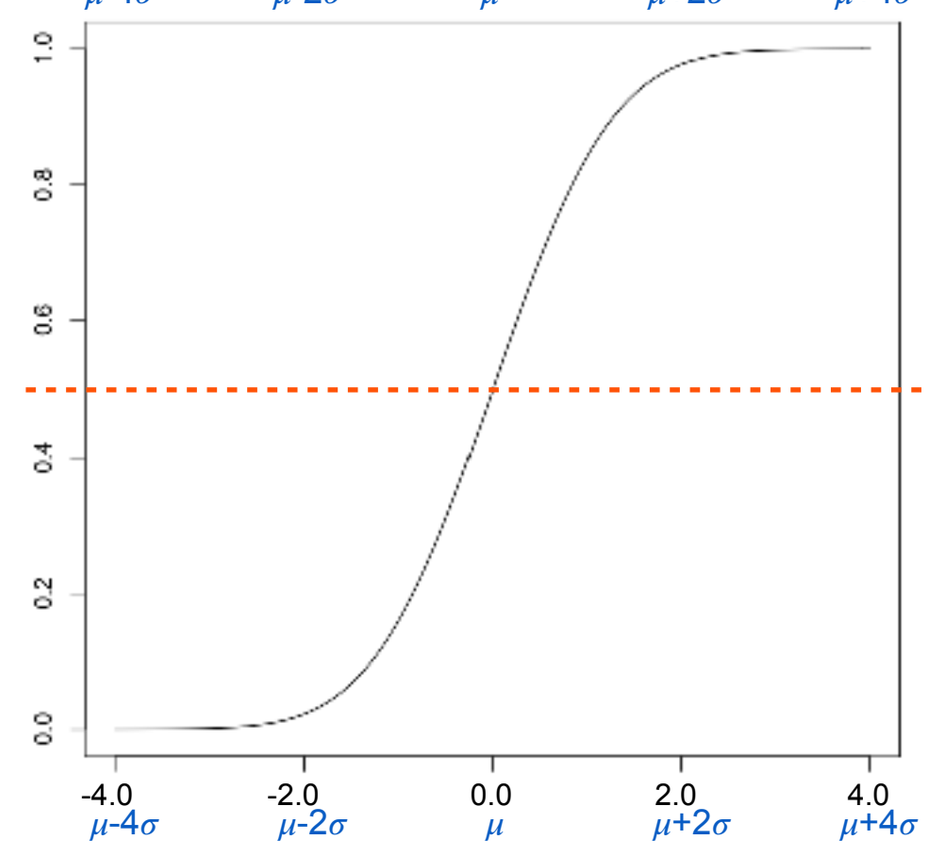
Symetrie: $X \sim N(0,1) \;\Rightarrow\; f_X(x) = f_X(-x)$

$Y \sim N(\mu, \sigma^2) \;\Rightarrow f_Y(\mu + x) = f_Y(\mu - x)$



$X \sim N(0,1) \;\Rightarrow\; \Phi(x) = 1 - \Phi(-x)$

$Y \sim N(\mu, \sigma^2) \;\Rightarrow\; F(\mu + x) = 1 - F(\mu - x)$



$\Phi^{-1}(1-p) = -\Phi^{-1}(p)$

$F^{-1}(1-p) = \mu + (\mu - F^{-1}(p)) = 2\mu - F^{-1}(p)$

7) Model normálního rozdělení

Příklady:

$$X \sim N(0,1), \quad Y \sim N(\mu, \sigma^2)$$

1) $P(X > x) \;\; = 1 - P(X \leq x) = 1 - \Phi(x)$

2) $P(|X| \leq a) \;\; = \Phi(a) - \Phi(-a) = \Phi(a) - (1 - \Phi(a)) = 2\Phi(a) - 1$

3) $P(|Y - \mu| \leq b) = F(\mu + b) - F(\mu - b) = F(\mu + b) - (1 - F(\mu + b)) = 2F(\mu + b) - 1$

7) Model normálního rozdělení

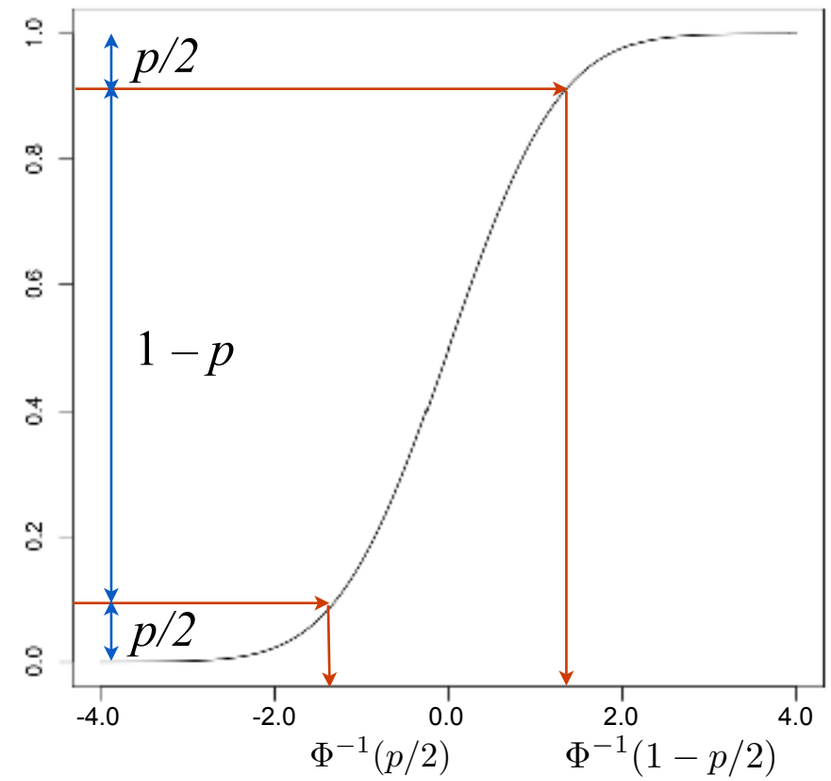
Interval spolehlivosti

Najděte symetrický interval, v němž budou hodnoty náhodné veličiny X s danou pravděpodobností p.

$X \sim N(0,1)$

Označme $\Phi^{-1}(1-p/2) = u_p$, potom ze symetrie máme $\Phi^{-1}(p/2) = -u_p$ (tzv. p-kritická hodnota)

Potom hledaný interval je $\langle -u_p; u_p \rangle$.



$Y \sim N(\mu, \sigma^2)$ => Najděte symetrický interval, v němž budou hodnoty náhodné veličiny Y s danou pravděpodobností p.

7) Model normálního rozdělení

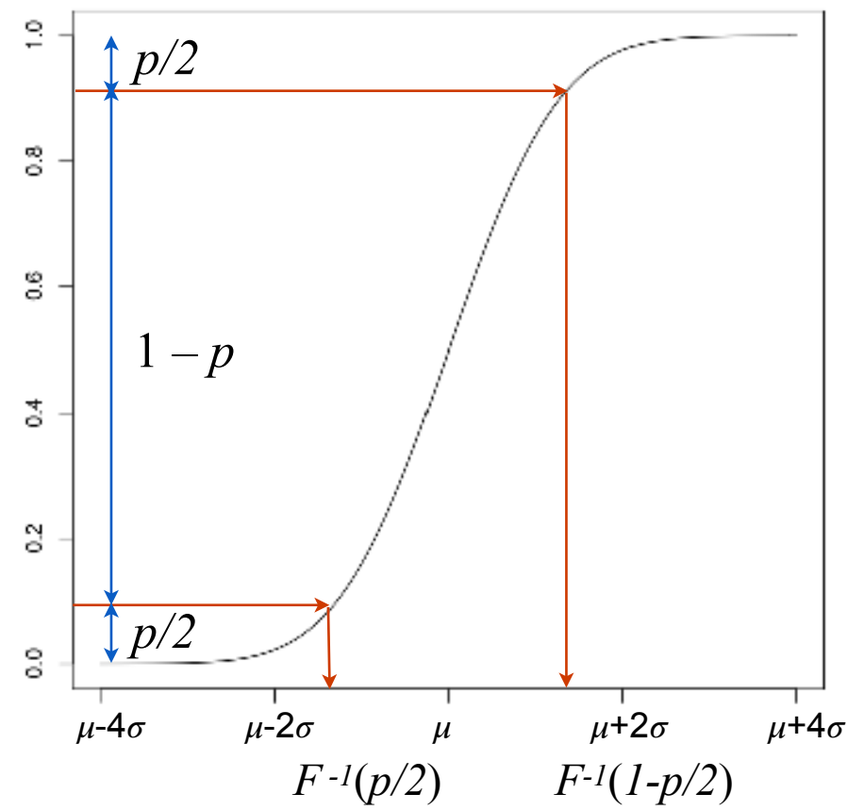
Interval spolehlivosti

Najděte symetrický interval, v němž budou hodnoty náhodné veličiny X s danou pravděpodobností p.

$X \sim N(0,1)$

Označme $\Phi^{-1}(1-p/2) = u_p$, potom ze symetrie máme $\Phi^{-1}(p/2) = -u_p$ (tzv. p-kritická hodnota)

Potom hledaný interval je $\langle -u_p; u_p \rangle$.



$Y \sim N(\mu, \sigma^2)$ => Najděte symetrický interval, v němž budou hodnoty náhodné veličiny Y s danou pravděpodobností p.

Označme nyní $F^{-1}(1-p/2) = r_p$, potom ze symetrie plyne $F^{-1}(p/2) = 2\mu - r_p$

a dále $P(Y \le r_p) = P(\sigma X + \mu \le r_p) = P\left(X \le \dfrac{r_p - \mu}{\sigma}\right) = 1 - p/2$.

Tedy $\dfrac{r_p - \mu}{\sigma} = u_p \quad \Rightarrow \quad r_p = \sigma u_p + \mu$. Potom hledaný interval je $\langle \mu - \sigma u_p ; \mu + \sigma u_p \rangle$.

8) Modely odvozené z normálního rozdělení

- **Model součtu normálně rozdělených náhodných veličin**

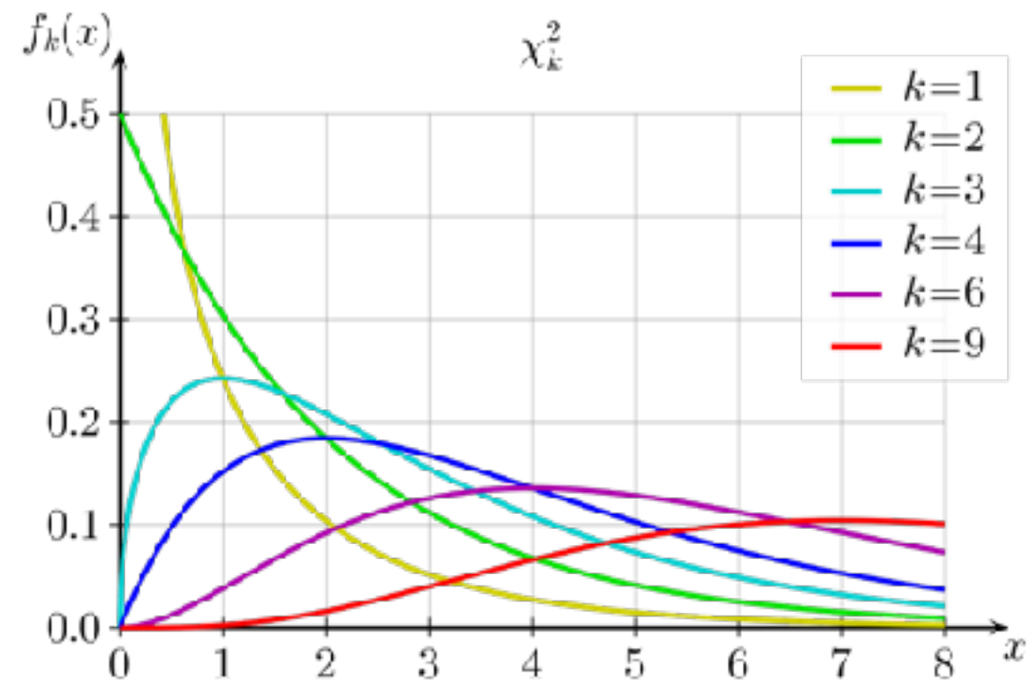
$$X \sim N(0,1) \quad \Rightarrow \quad X^2 \sim \chi^2(1)$$

$$X_i \sim N(0,1) \quad i = 1, 2, \ldots, k$$ + nezávislost

$$\Rightarrow \quad S_k^2 = \sum_{i=1}^{k} X_i^2 \sim \chi^2(k)$$

Pearsonovo (χ^2) rozdělení s k stupni volnosti

$$E(S_k^2) = k, \quad Var(S_k^2) = 2k$$



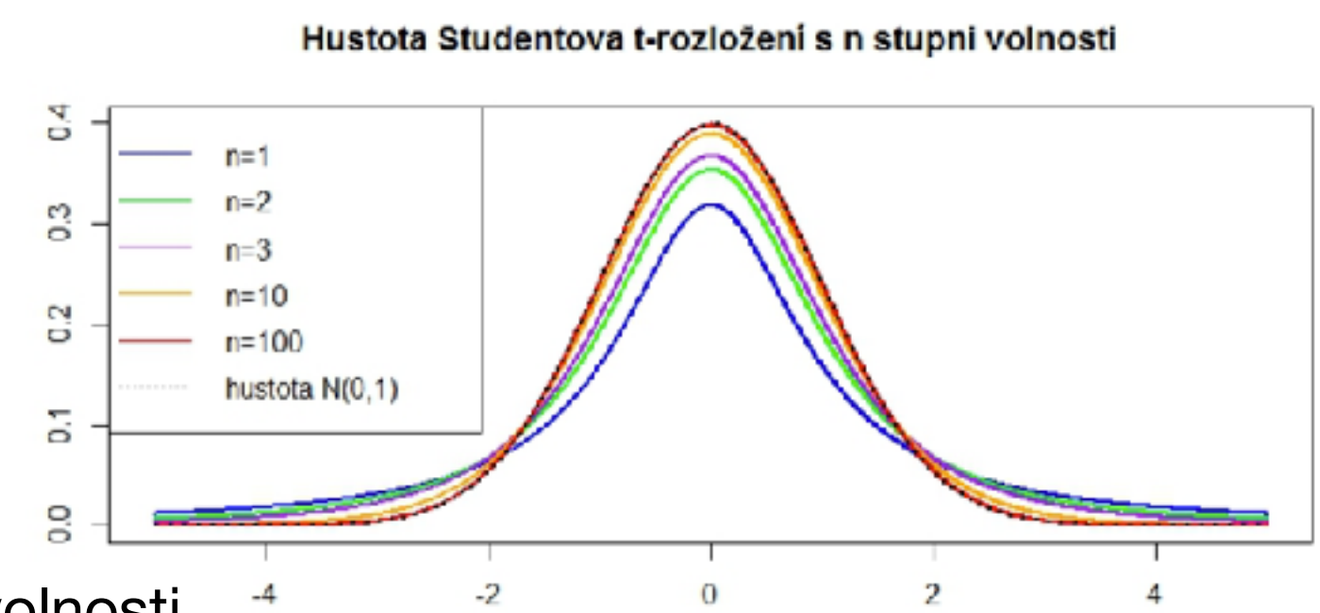
- **Model podílu normálně rozdělené náhodné veličiny a veličiny s rozdělením chí-kvadrát**

Jsou-li U a V stoch. nezávislé n.v.:

$$U \sim N(0,1), \quad V \sim \chi^2(n),$$

$$\Rightarrow \quad T = \frac{U}{\sqrt{\frac{V}{n}}} \sim t(n)$$

Studentovo t-rozdělení o n stupních volnosti



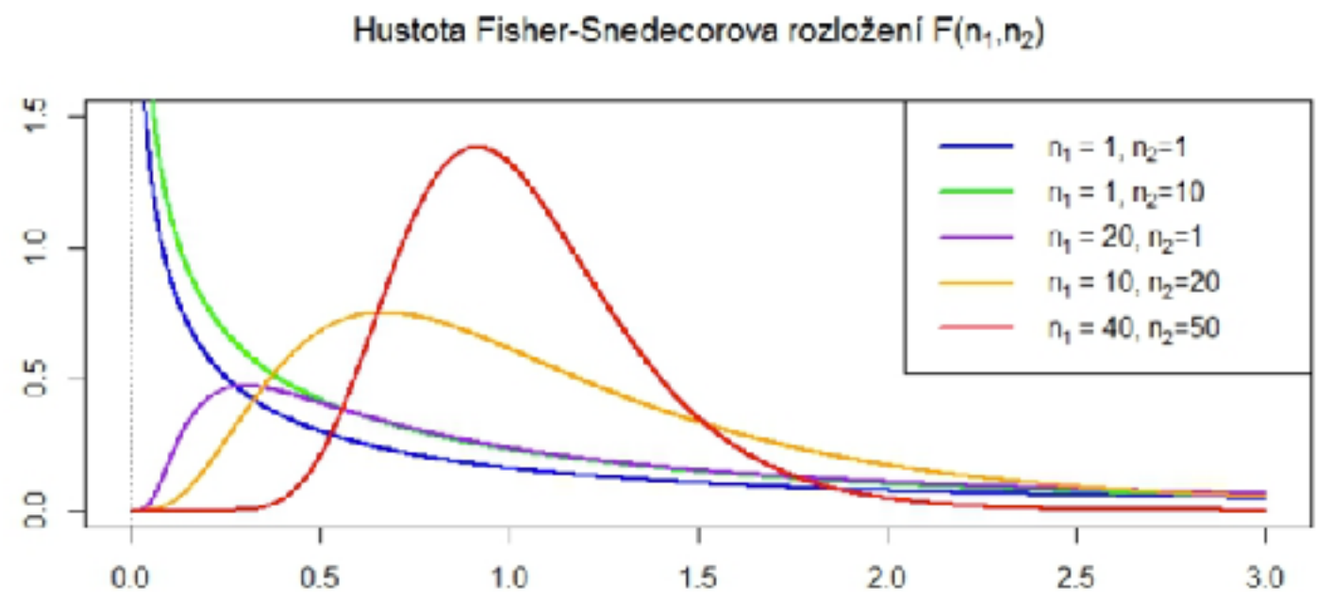
8) Modely odvozené z normálního rozdělení

- **Model podílu dvou chí-kvadrát rozdělených veličin**

$$X_i, Y_j \sim \ N(0,1) \ \text{ i.i.d.}$$

$$R_{m,n} = \frac{\frac{1}{m}\sum_{i=1}^{m} X_i^2}{\frac{1}{n}\sum_{j=1}^{n} Y_j^2} = \frac{nS_m^2}{mS_n^2} \sim \ F(m,n)$$

Fischerovo-Snedecorovo rozdělení s m a n stupni volnosti.



- **Model logaritmicko-normálního rozdělení**

$$Y \sim \ N(\mu, \sigma^2), \quad X = \exp(Y)$$

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad x > 0$$

$$E(X) = e^{\mu + \sigma^2/2}, \quad Var(X) = \left(e^{\sigma^2} - 1\right) e^{2\mu + \sigma^2}$$

Logaritmicko-normální rozdělení $\mathrm{LN}(\mu, \sigma^2)$