Základy pravděpodobnosti a matematické statistiky

5. Model s normálním rozdělením

FAKULTA STROJNÍ ČVUT V PRAZE

prof. RNDr. Gejza Dohnal, CSc.

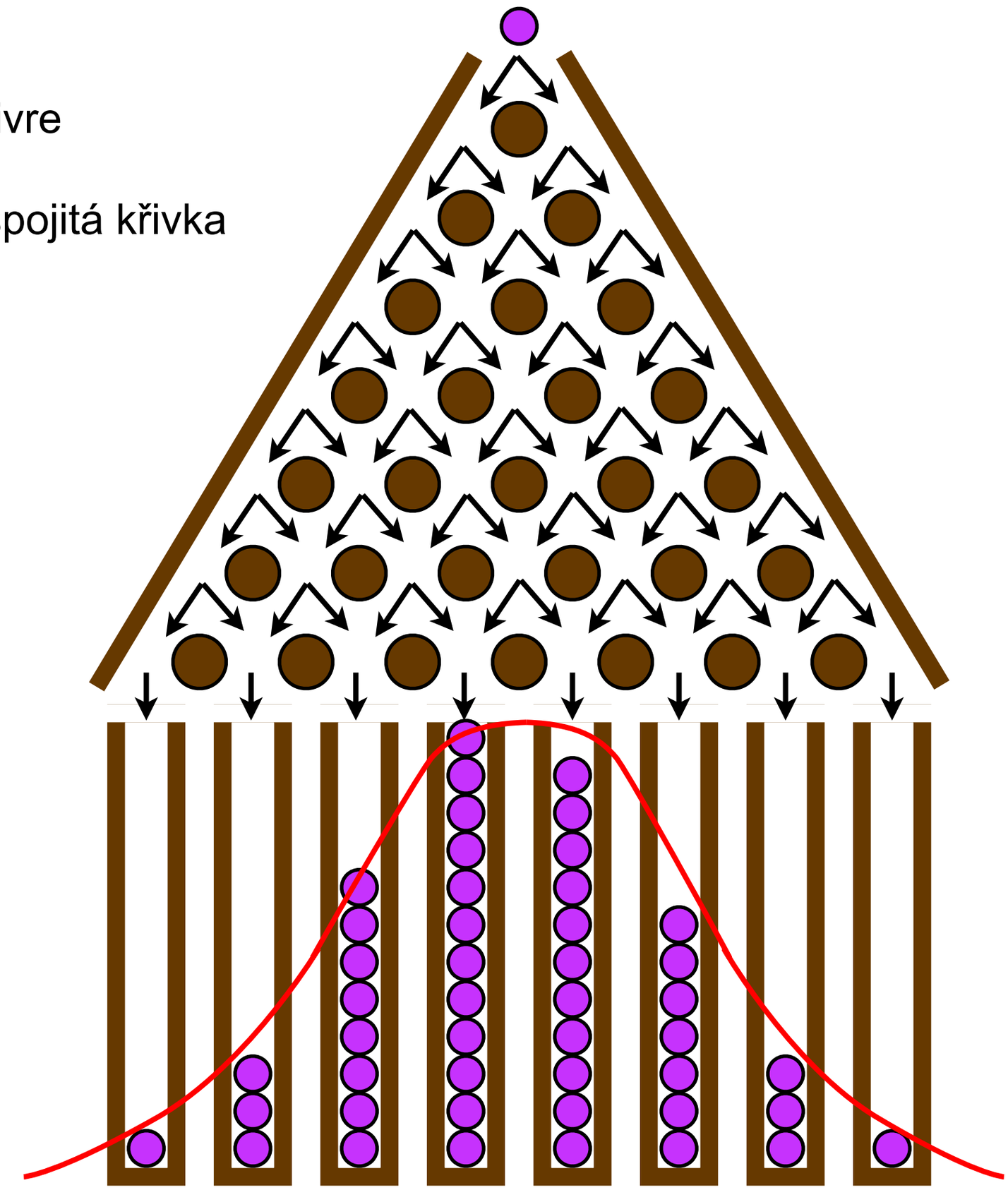
ak. rok 2023/2024

5. Model s normálním rozdělením

Galtonova deska

1733 : Abraham de Moivre

pokusy s mincemi ⟶ spojitá křivka



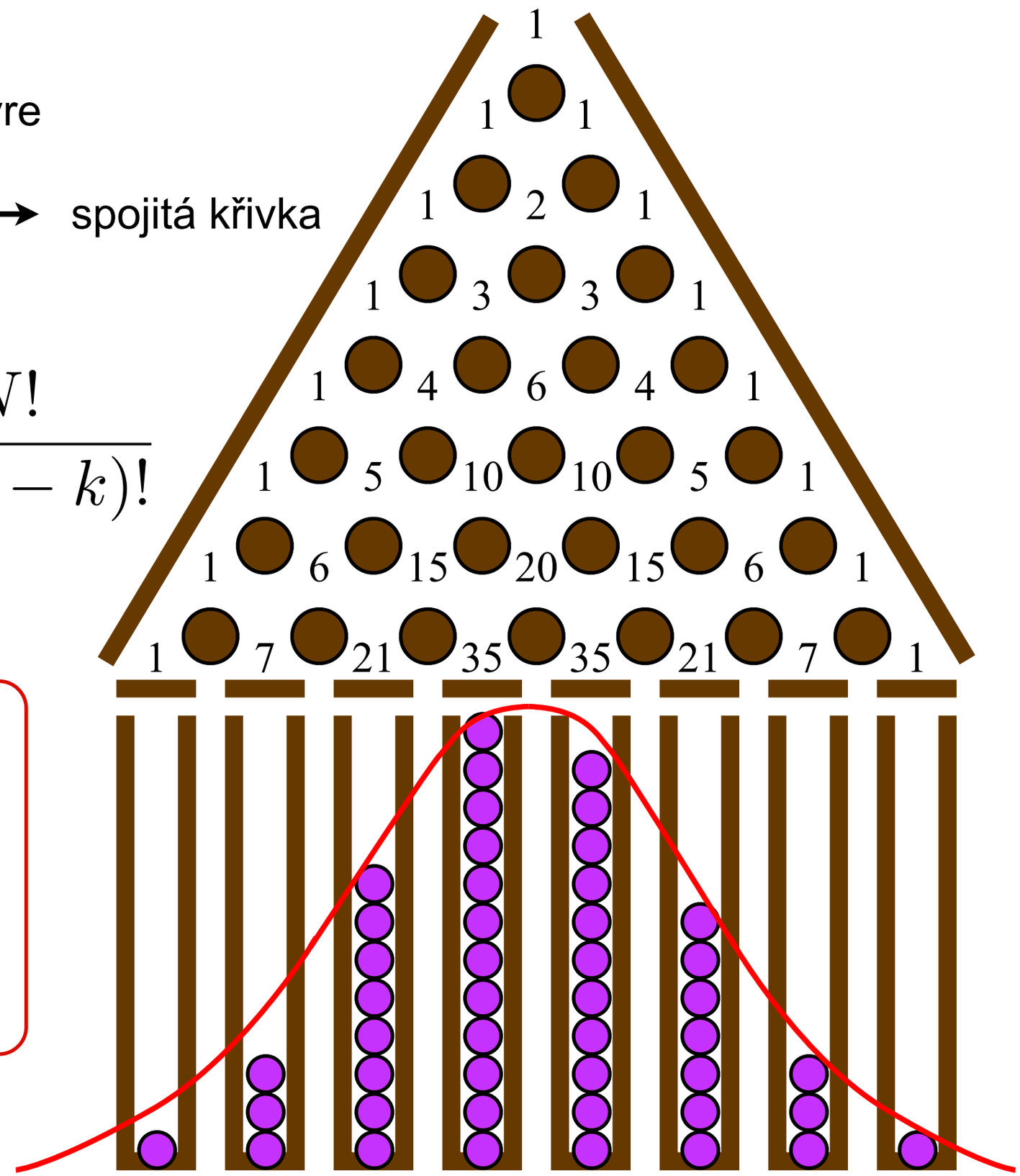
Galtonova deska

1733 : Abraham de Moivre

pokusy s mincemi $\longrightarrow$ spojitá křivka

Pascalův trojúhelník

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



Moivre navrhl rovnici:

$$y = e^{-\frac{x^2}{2}}$$

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

Rozdělení chyb měření

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

Karl Friedrich Gauss a Piere Laplace: “křivka chyb měření”

gaussova křivka

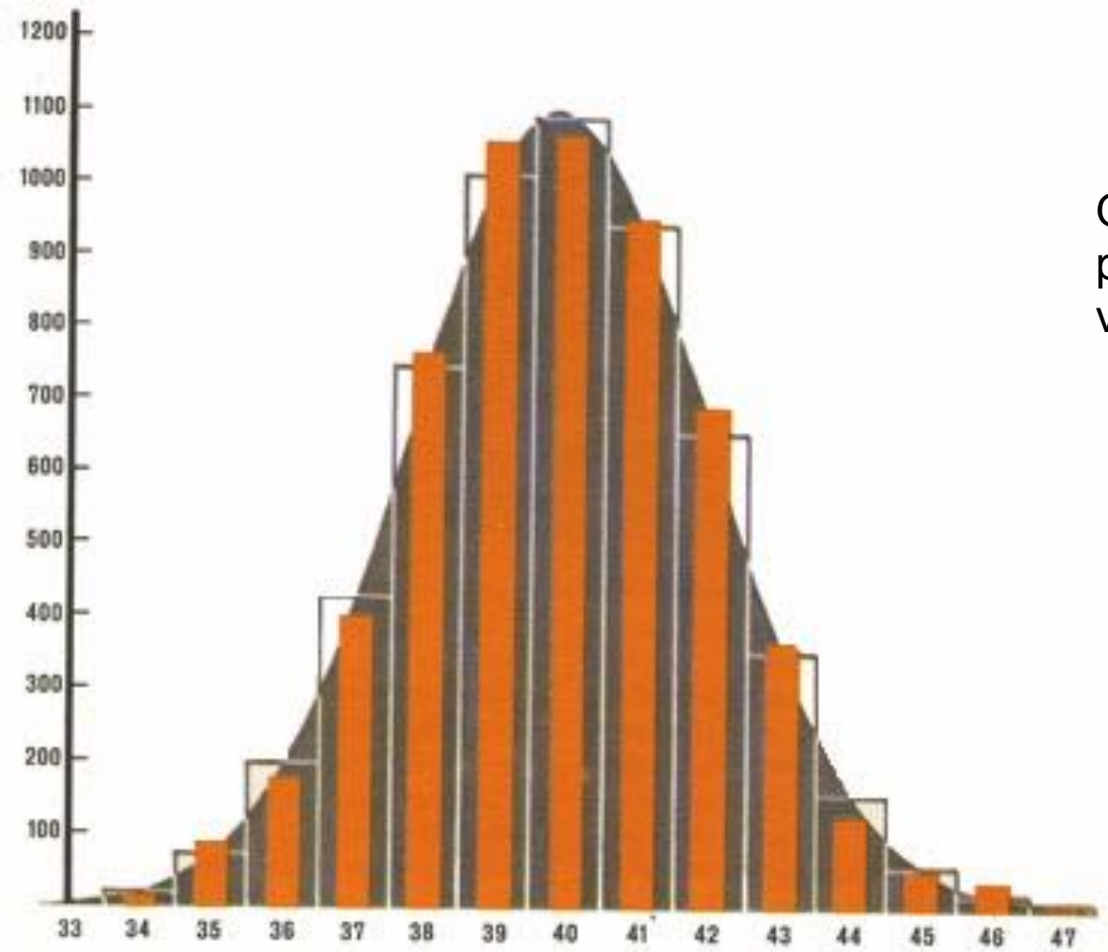
Model „normálního“ vojáka

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

Karl Friedrich Gauss a Piere Laplace: “křivka chyb měření”

gaussova křivka

počátek 19. století: Adolphe Qételet
- antropometrická měření anglických vojáků



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

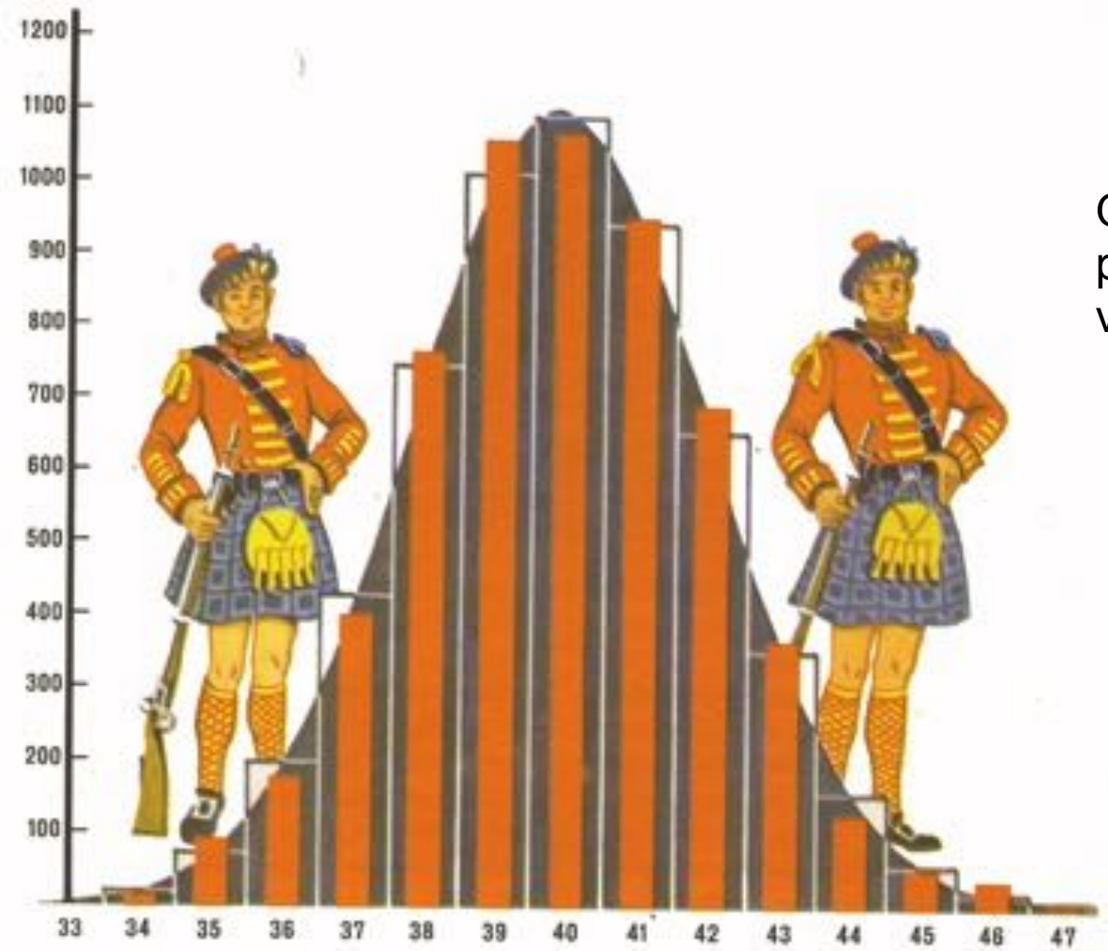
Normální (Gaussovo) rozdělení

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

Karl Friedrich Gauss a Piere Laplace: “křivka chyb měření”

gaussova křivka

počátek 19. století: Adolphe Qételet
- antropometrická měření anglických vojáků



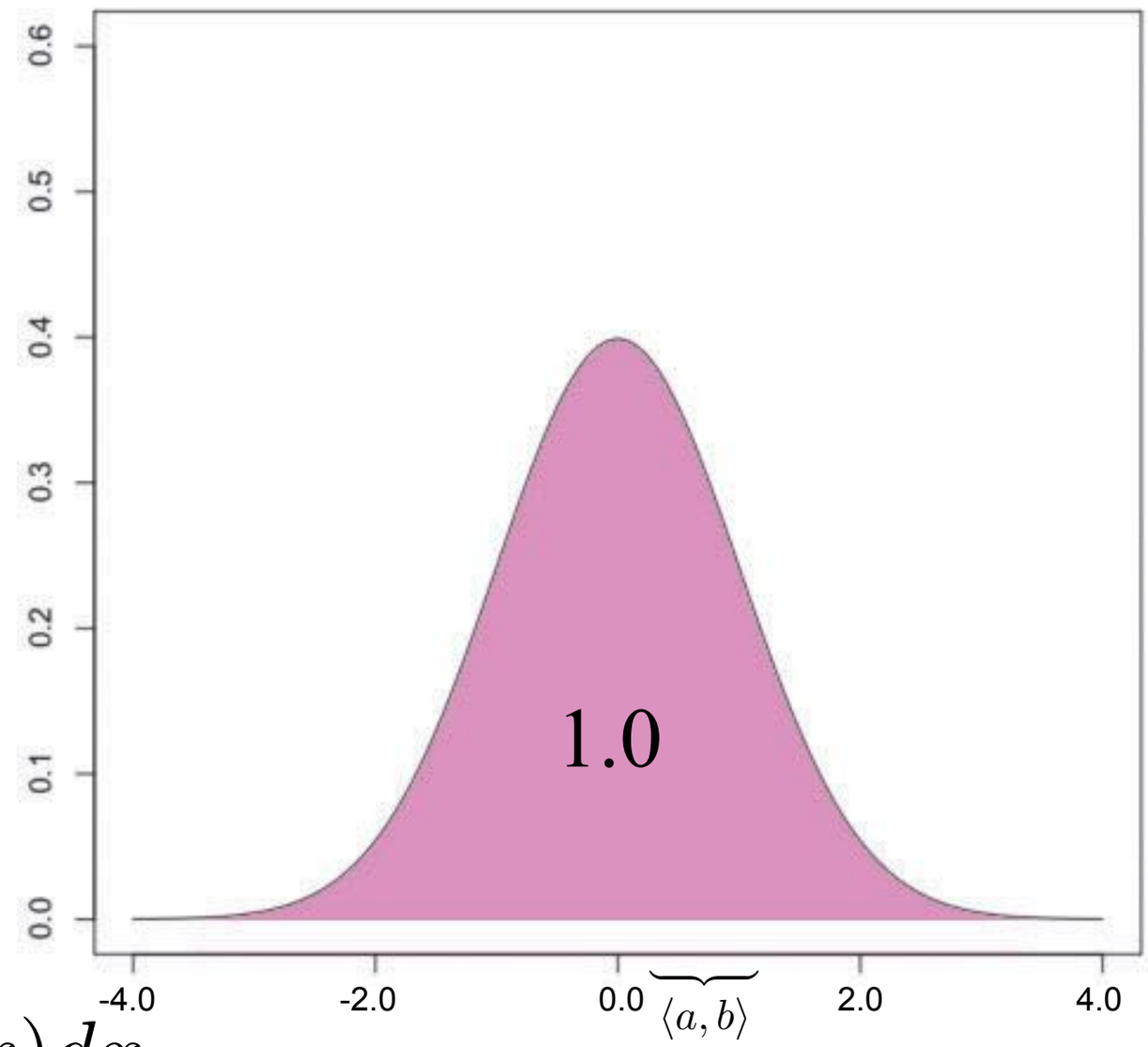
Obvod hrudi skotských vojáků v palcích, naměřený Quételetem ve dvacátých letech 19. století

normální rozdělení

Normální (Gaussovo) rozdělení

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

$$x \in R$$



Standardní normální rozdělení $N(0,1)$

$$X \sim N(0,1)$$

$$X \in (-\infty, +\infty)$$

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

Standardizace

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

Normální (Gaussovo) rozdělení

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

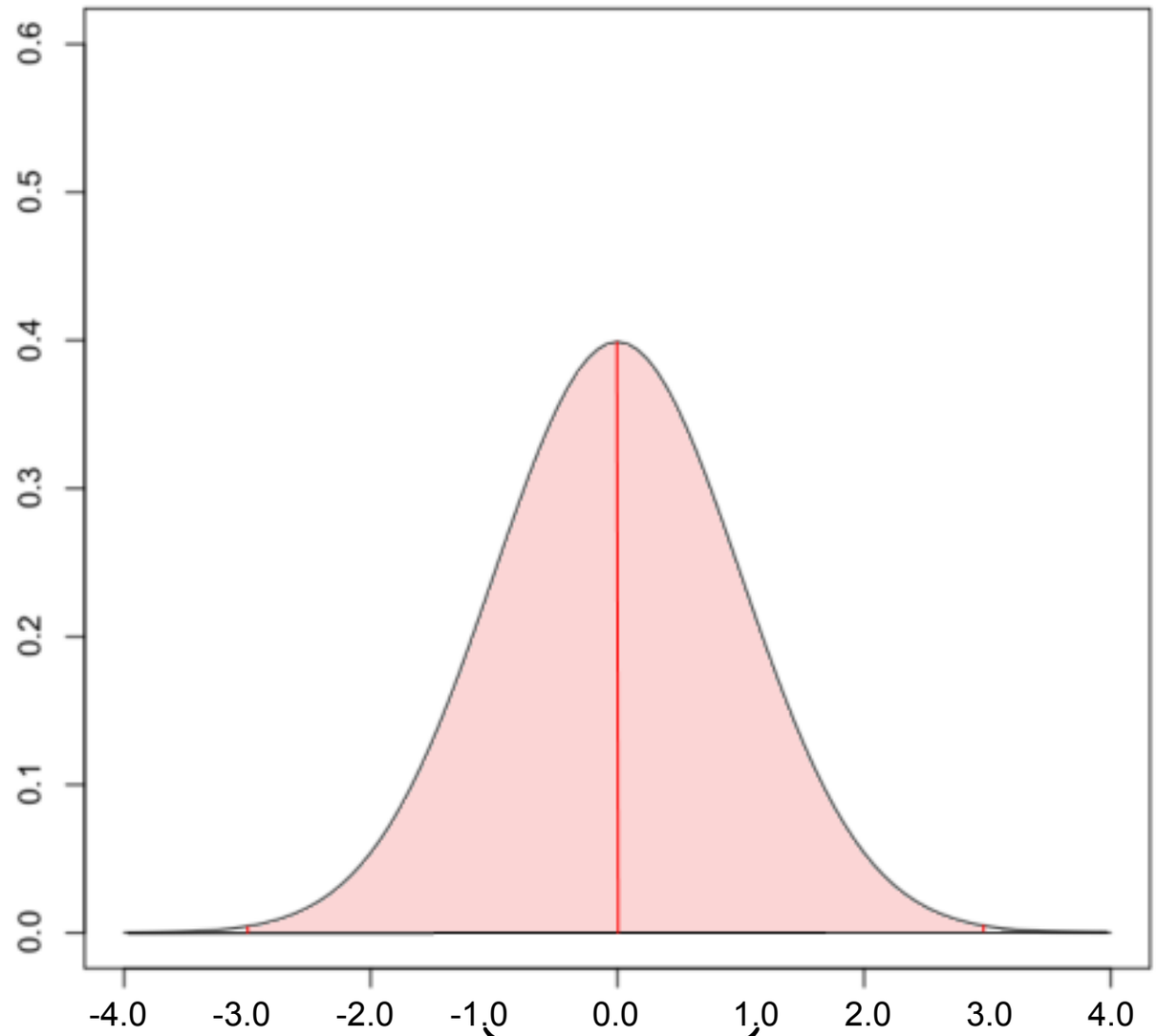
$$x \in R$$

$$X \sim N(0,1)$$

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

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

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



$$Y \sim N(\mu, \sigma^2)$$

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

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

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

$\langle \mu - \sigma, \mu + \sigma \rangle$
68,26%

$\langle \mu - 2\sigma, \mu + 2\sigma \rangle$
95%

$\langle \mu - 3\sigma, \mu + 3\sigma \rangle$
99.7%

Normální (Gaussovo) rozdělení

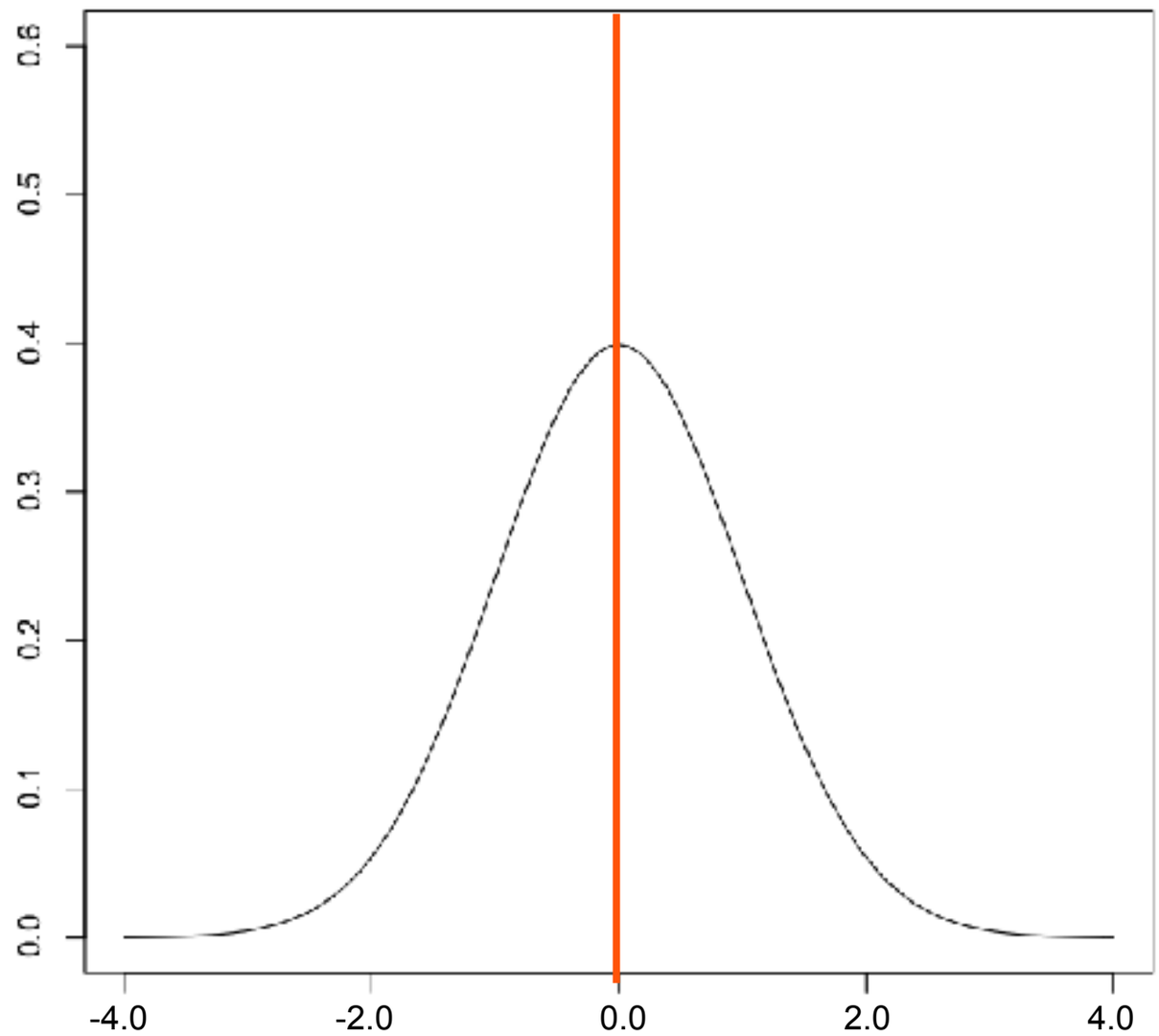
Symetrie:

$$X \sim N(0,1)$$

$$f(x) = f(-x)$$

$$X \sim N(\mu, \sigma^2)$$

$$f(\mu + x) = f(\mu - x)$$



Normální (Gaussovo) rozdělení

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

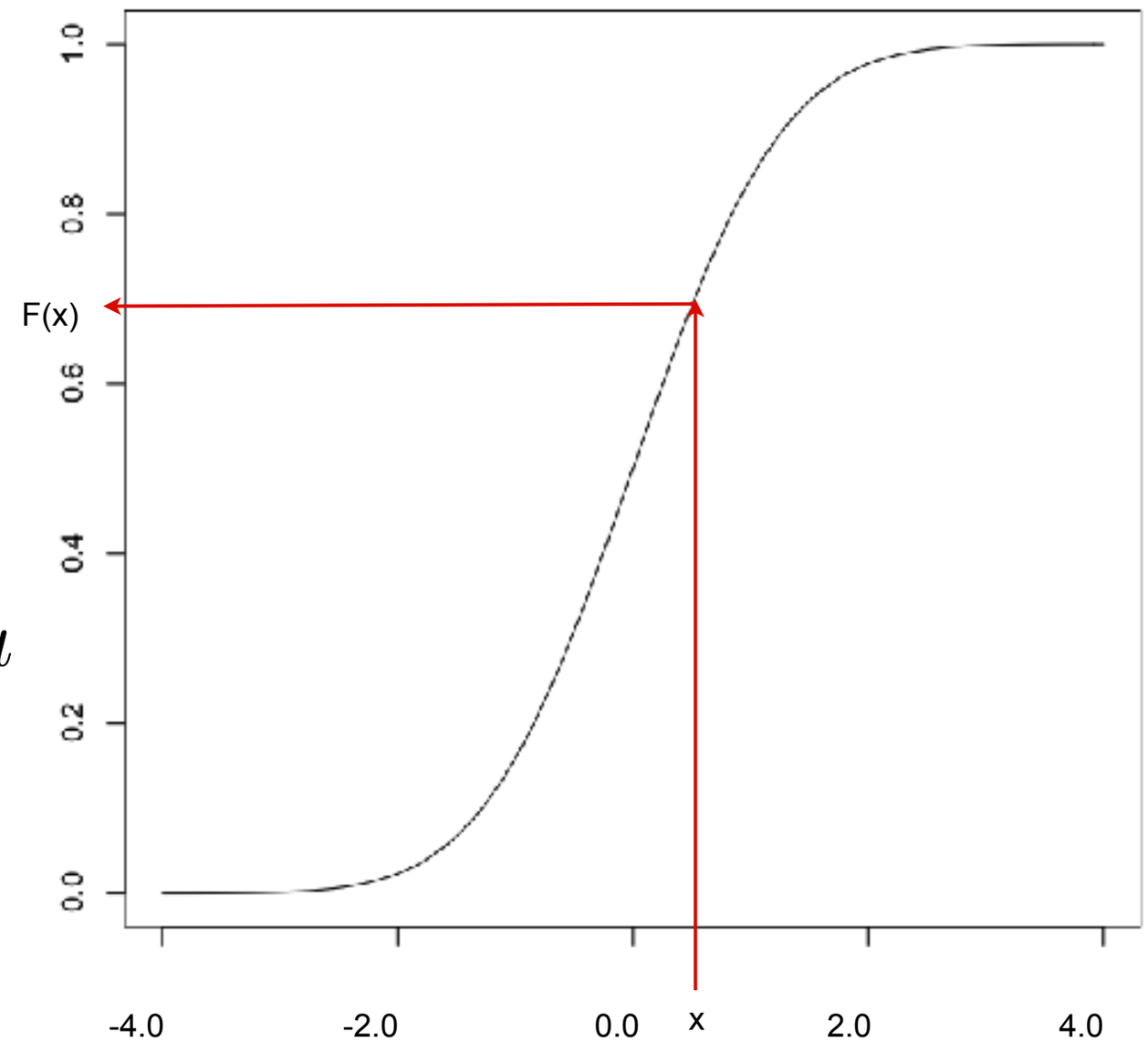
$$x \in R$$

$$X \sim N(0,1)$$

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

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

distribuční funkce N(0,1)



S jakou pravděpodobností bude hodnota náhodné veličiny X menší než x?

S jakou pravděpodobností nebude hodnota náhodné veličiny X větší než x?

S jakou pravděpodobností bude hodnota náhodné veličiny X větší než x?

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

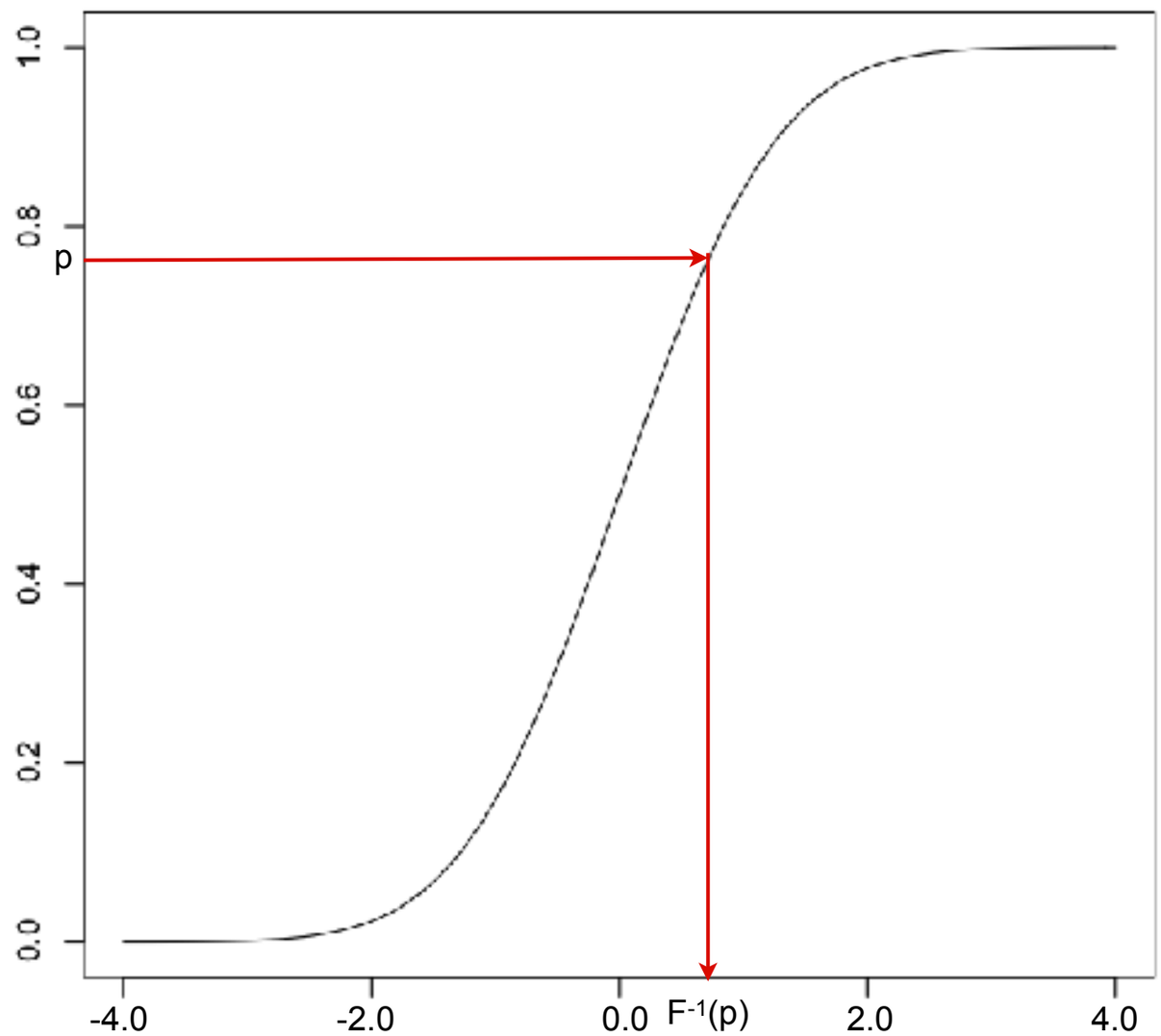
Normální (Gaussovo) rozdělení

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

$$x \in R$$

$$X \sim N(0,1)$$

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



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

distribuční funkce

Jakou hodnotu náhodná veličina X nepřekročí s pravděpodobností p?

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

kvantilová funkce

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

Normální (Gaussovo) rozdělení

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

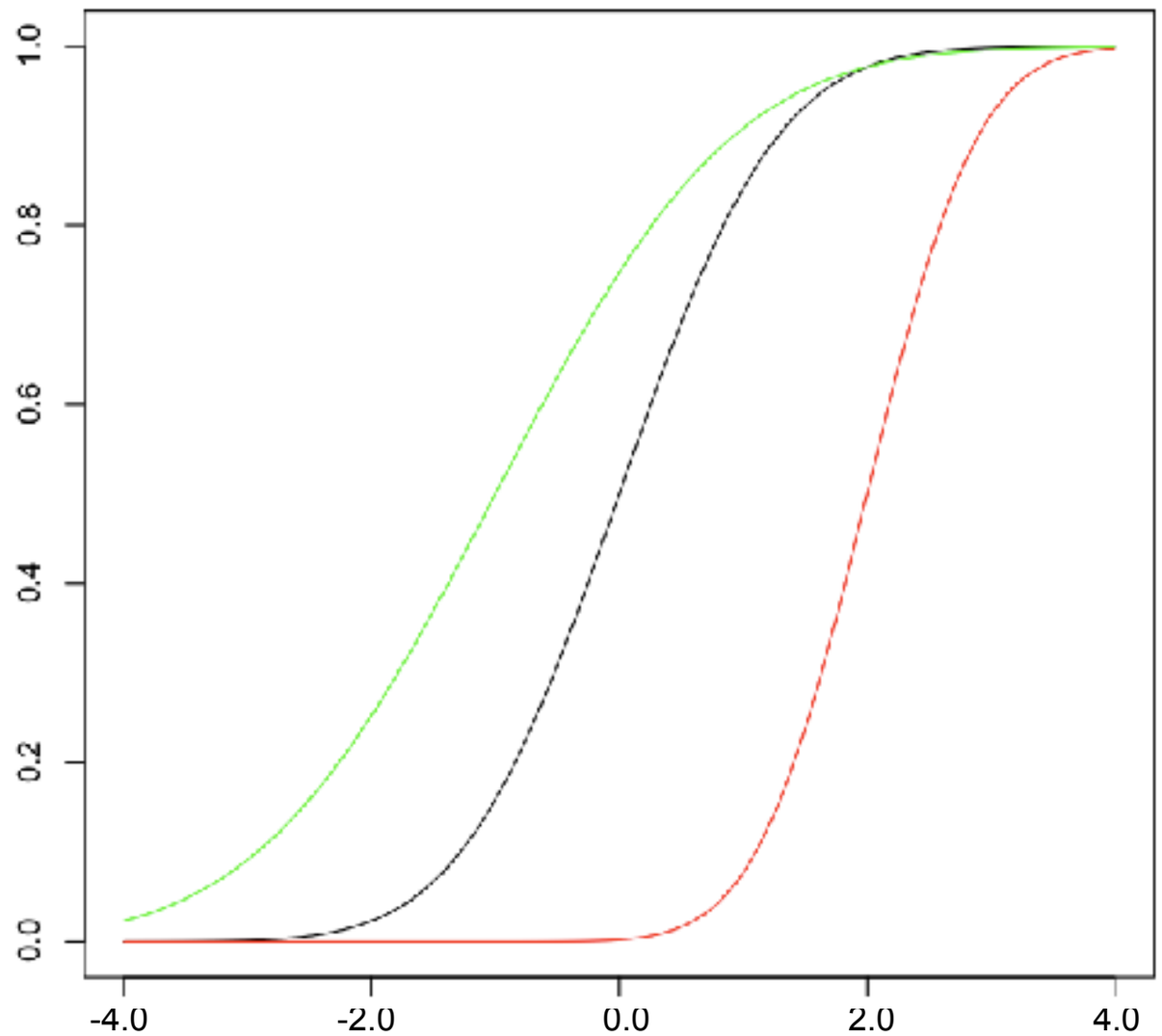
$$x \in R$$

$$X \sim N(\mu, \sigma^2)$$

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

$$\boxed{F(x) = P(X \leq x)}$$

distribuční funkce



$$F(x) = \int_{-\infty}^{x} f(u)du$$

$$f(x) = \frac{dF(x)}{dx}, \quad x \in R$$

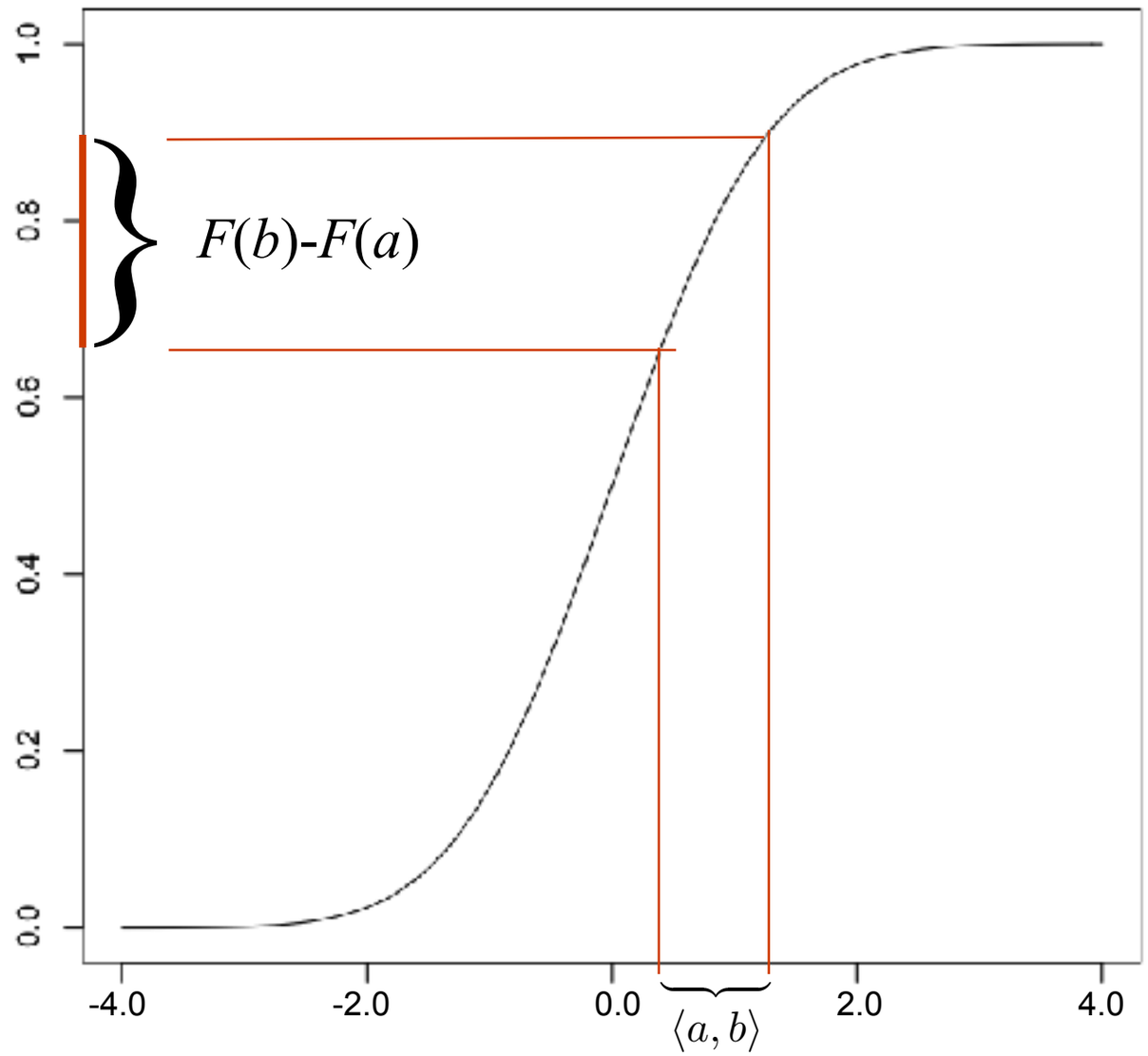
Normální (Gaussovo) rozdělení

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

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

$$= \int_{-\infty}^b f(x)dx - \int_{-\infty}^a f(x)dx$$

$$\boxed{P(a \leq X \leq b) = F(b) - F(a)}$$



Normální (Gaussovo) rozdělení

Symetrie:

$$X \sim N(0,1)$$

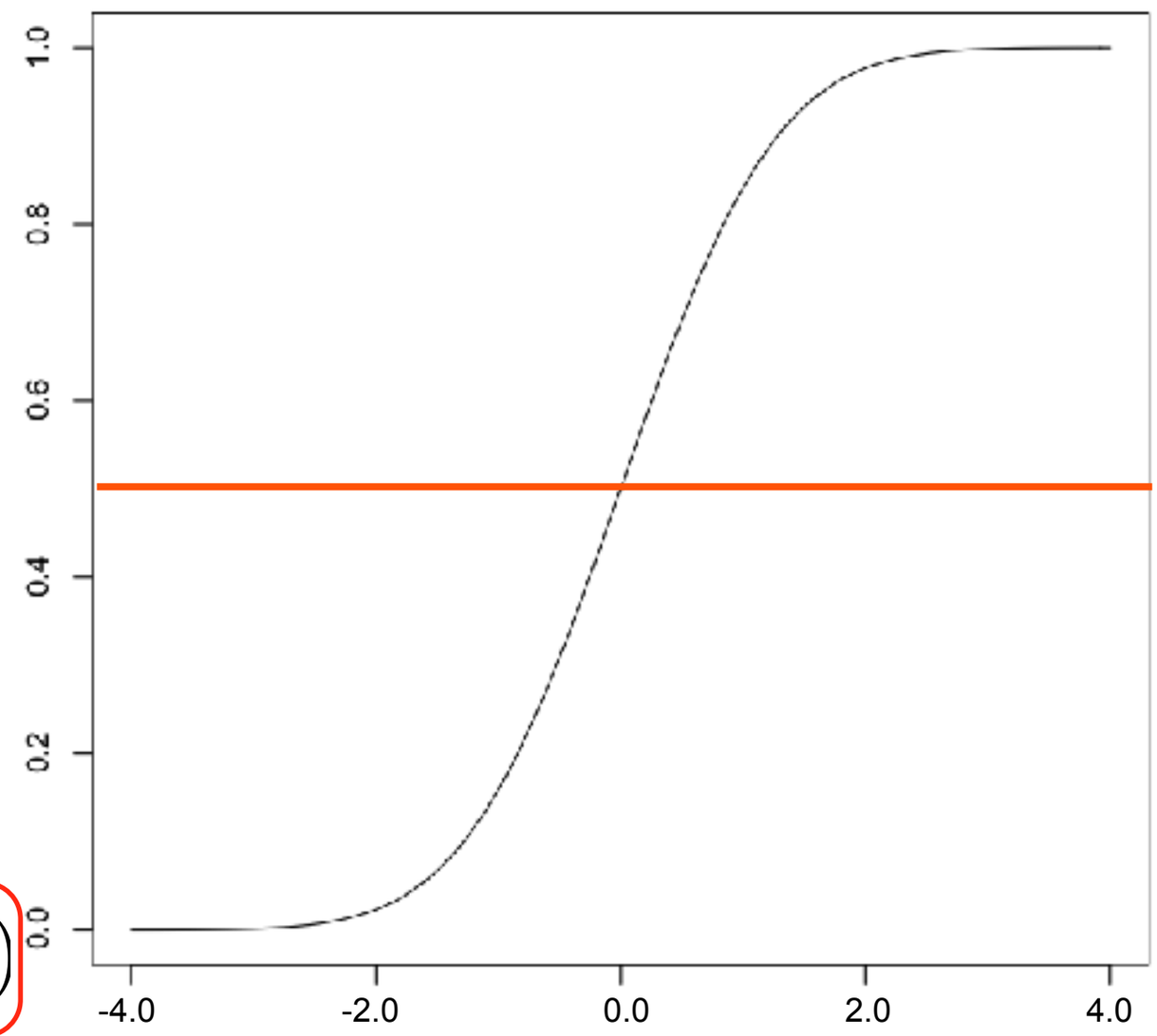
$$f(x) = f(-x)$$

$$\boxed{\Phi(x) = 1 - \Phi(-x)}$$

$$X \sim N(\mu, \sigma^2)$$

$$f(\mu + x) = f(\mu - x)$$

$$\boxed{F(\mu + x) = 1 - F(\mu - x)}$$



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

$$P(|X - \mu| \leq x) = F(\mu + x) - F(\mu - x) =$$

$$F(\mu + x) - (1 - F(\mu + x)) = 2F(\mu + x) - 1$$

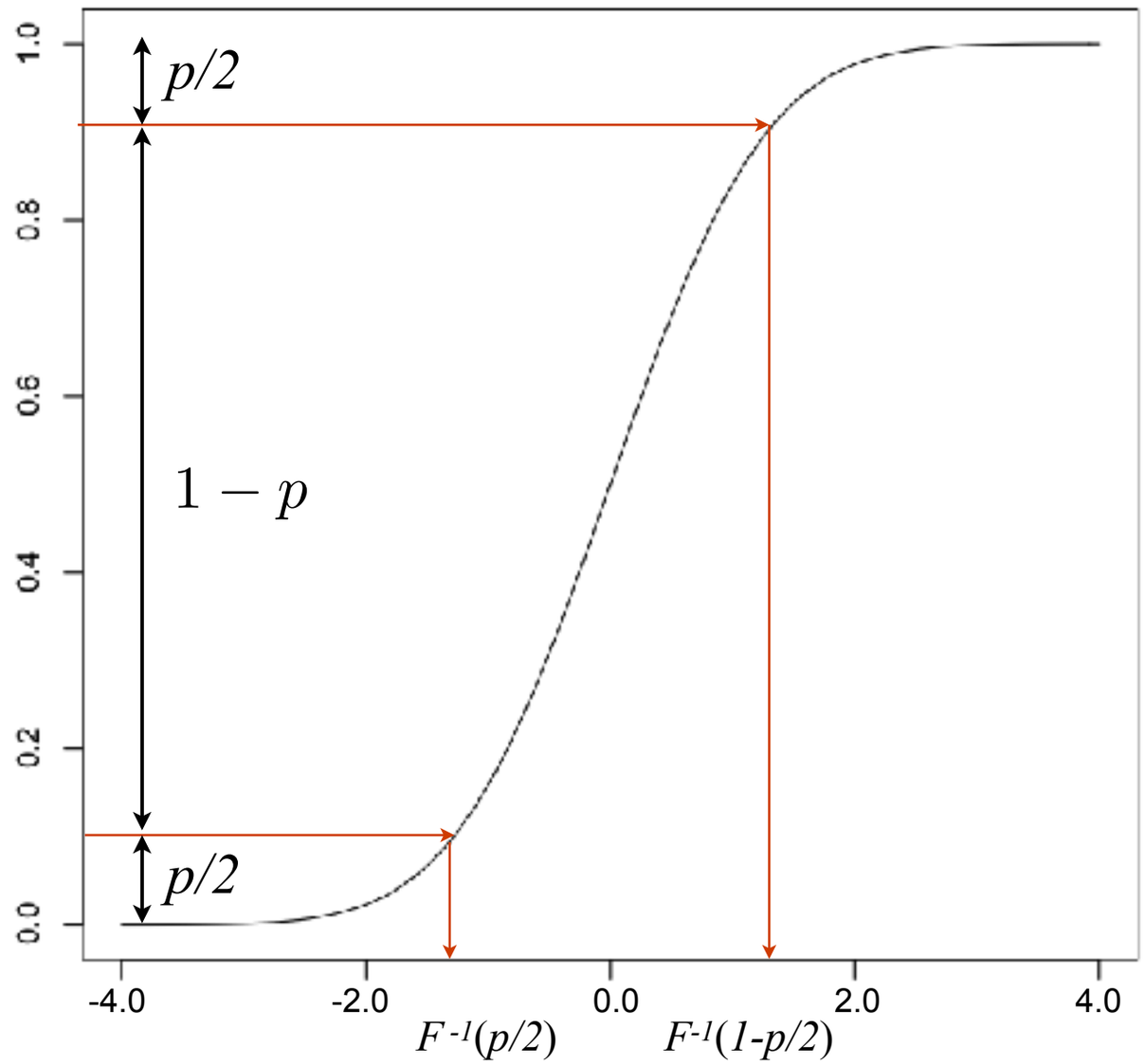
Normální (Gaussovo) rozdělení

Jakých hodnot kolem střední hodnoty bude náhodná veličina X nabývat s předem danou pravděpodobností 1-p?

$$P(a \leq X \leq b) = 1 - p$$

$$\begin{aligned} a &= F^{-1}(p/2), \\ b &= F^{-1}(1 - p/2) \end{aligned}$$

($p/2$)-kvantil normálního rozdělení a (1-$p/2$)-kvantil normálního rozdělení



Ze symetrie plyne:

$$X \sim N(0,1): \quad x_{p/2} = \Phi^{-1}(p/2) = -x_{1-p/2} = -\Phi^{-1}(1 - p/2)$$

$$Y \sim N(\mu, \sigma^2): \quad y_{p/2} = F^{-1}(p/2) = 2\mu - y_{1-p/2} = 2\mu - F^{-1}(1 - p/2)$$

Normální (Gaussovo) rozdělení

Jaký je vztah mezi $x_{p/2}$ a $y_{p/2}$?

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

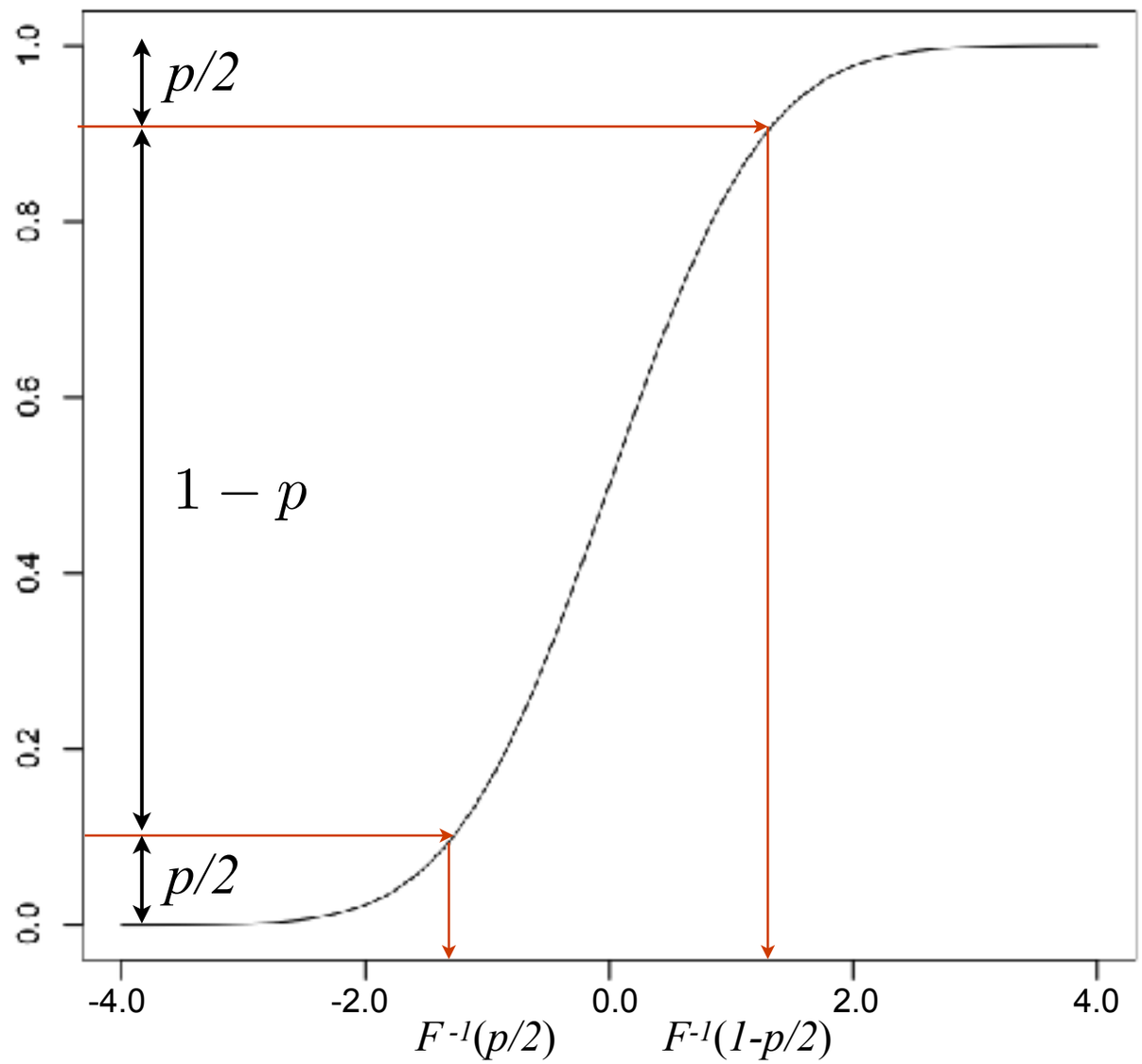
$$\Rightarrow X = \frac{Y-\mu}{\sigma}$$

$$\Rightarrow Y = \sigma X + \mu$$

$$F(y) = P(Y \leq y) =$$

$$P(\frac{Y-\mu}{\sigma} \leq \frac{y-\mu}{\sigma}) =$$

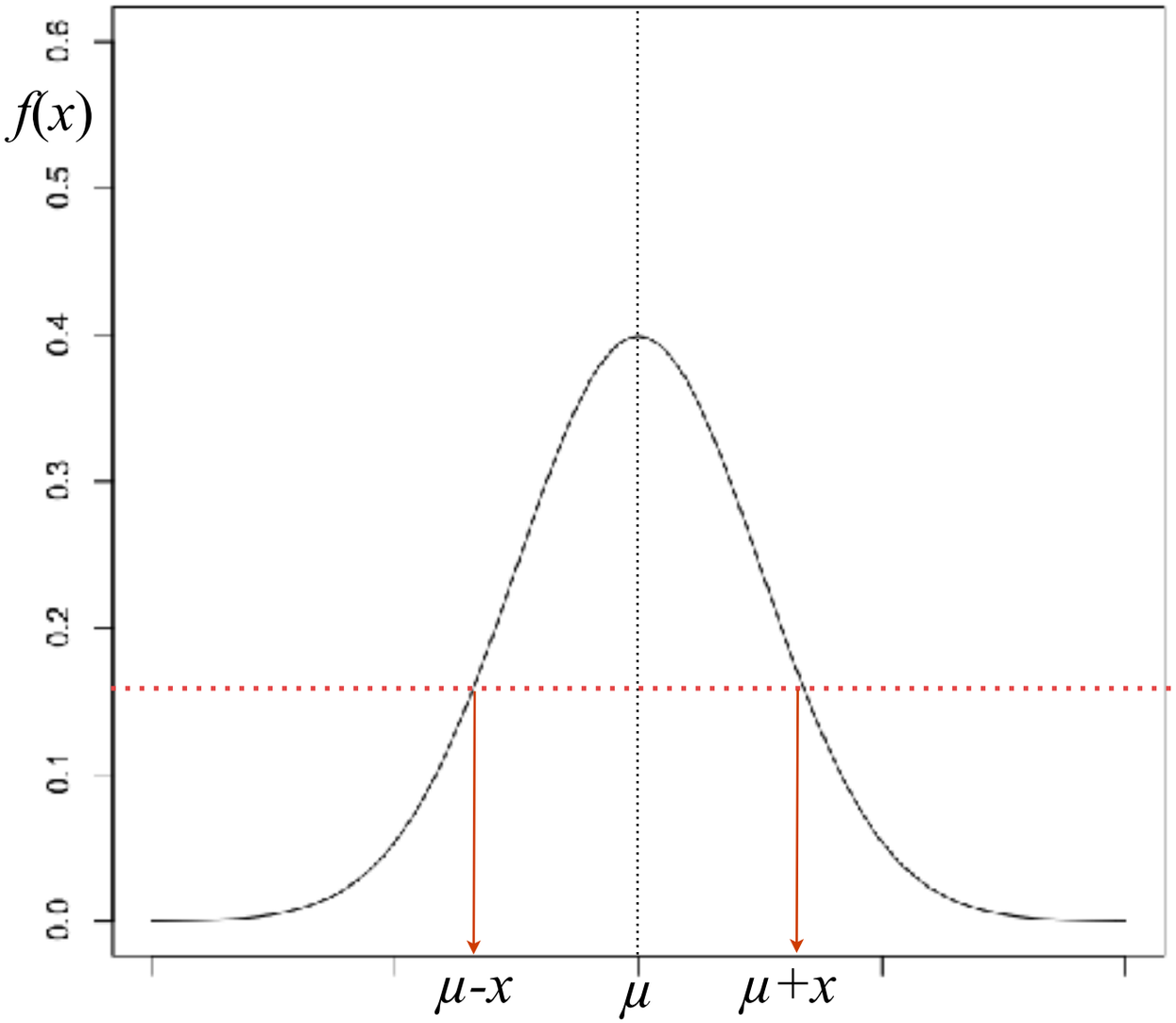
$$P(X \leq \frac{y-\mu}{\sigma}) = \Phi(\frac{y-\mu}{\sigma})$$

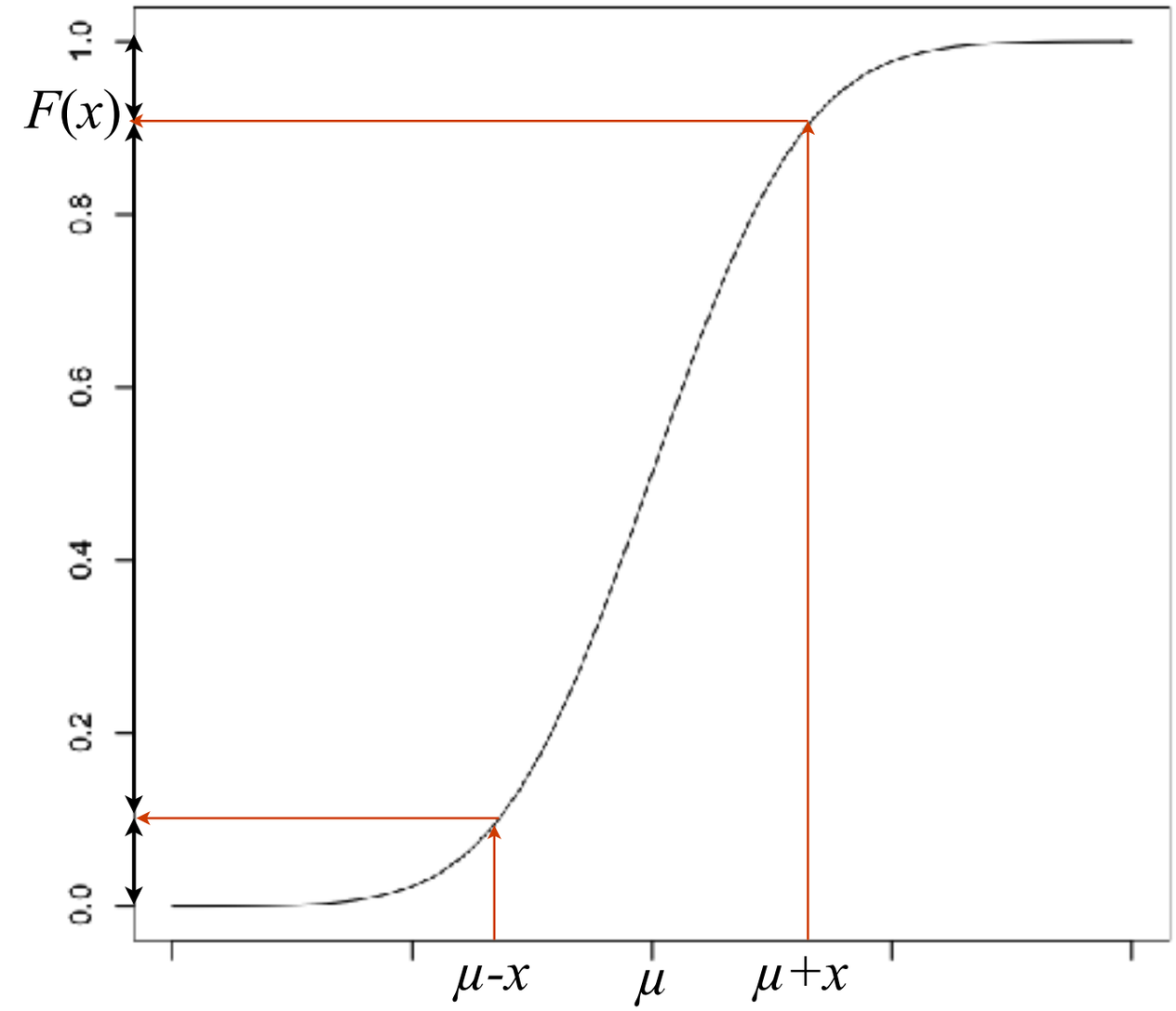


Tedy je: $$F(y) = \Phi(\frac{y-\mu}{\sigma})$$

... a odtud: $$F^{-1}(p) = \sigma\Phi^{-1}(p) + \mu$$

Vše se prakticky vyjadřuje pouze v tzv. kritických hodnotách: $u_p = \Phi^{-1}(1-p/2)$


f(x)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
μ-x
μ
μ+x


F(x)
0.0
0.2
0.4
0.6
0.8
1.0
μ-x
μ
μ+x

Normální (Gaussovo) rozdělení

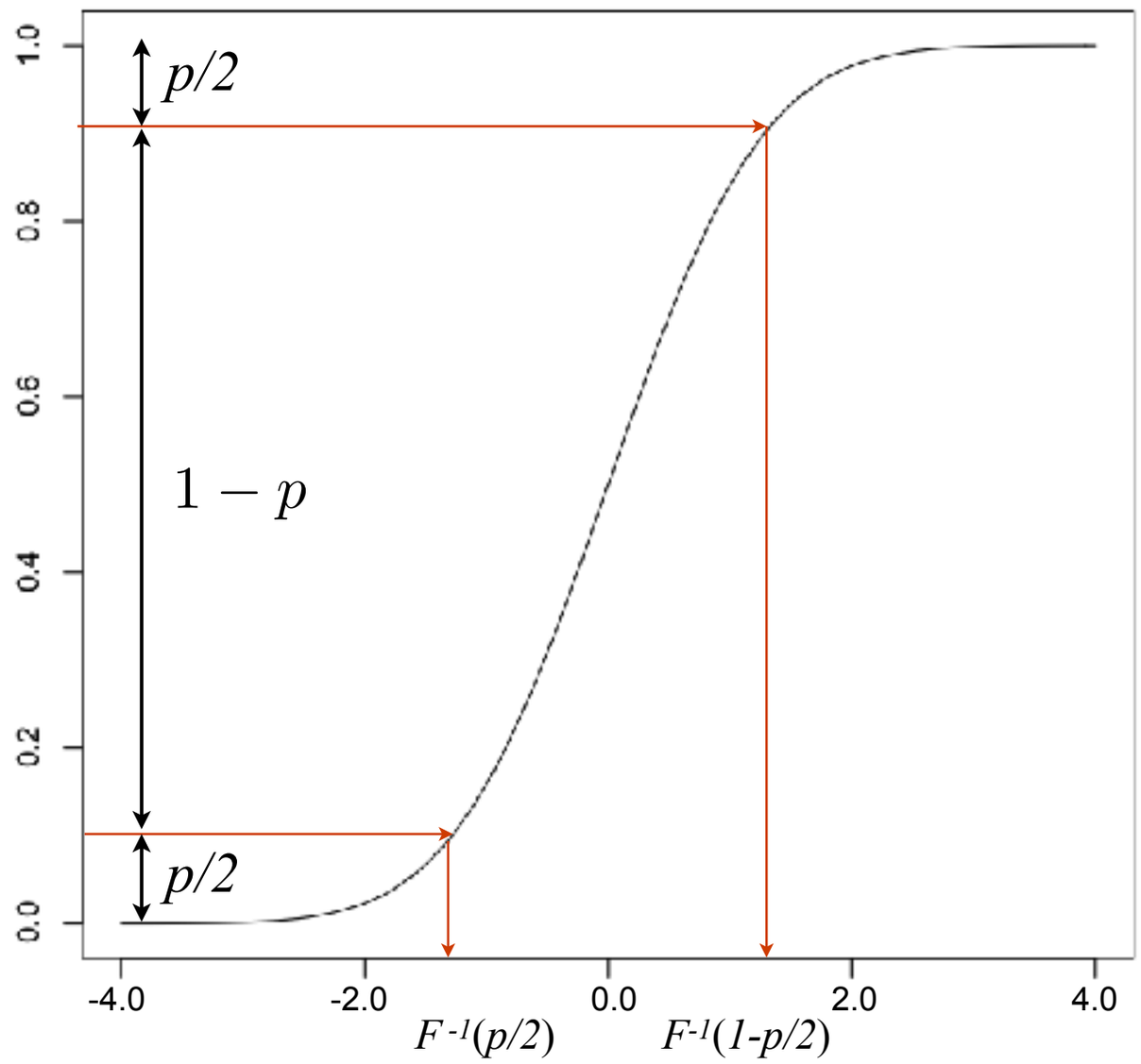
Jaký je vztah mezi $x_{p/2}$ a $y_{p/2}$?

... označme tedy

$$\Phi^{-1}(1-p/2) = u_p$$

... potom

$$\begin{aligned} \Phi^{-1}(p/2) &= -u_p \\ F^{-1}(1-p/2) &= \mu + \sigma u_p \\ F^{-1}(p/2) &= \mu - \sigma u_p \end{aligned}$$



Tedy je: $$F(y) = \Phi\left(\frac{y-\mu}{\sigma}\right)$$

... a odtud: $$F^{-1}(p) = \sigma\Phi^{-1}(p) + \mu$$

Vše se prakticky vyjadřuje pouze v tzv. kritických hodnotách: $u_p = \Phi^{-1}(1-p/2)$

Normální (Gaussovo) rozdělení

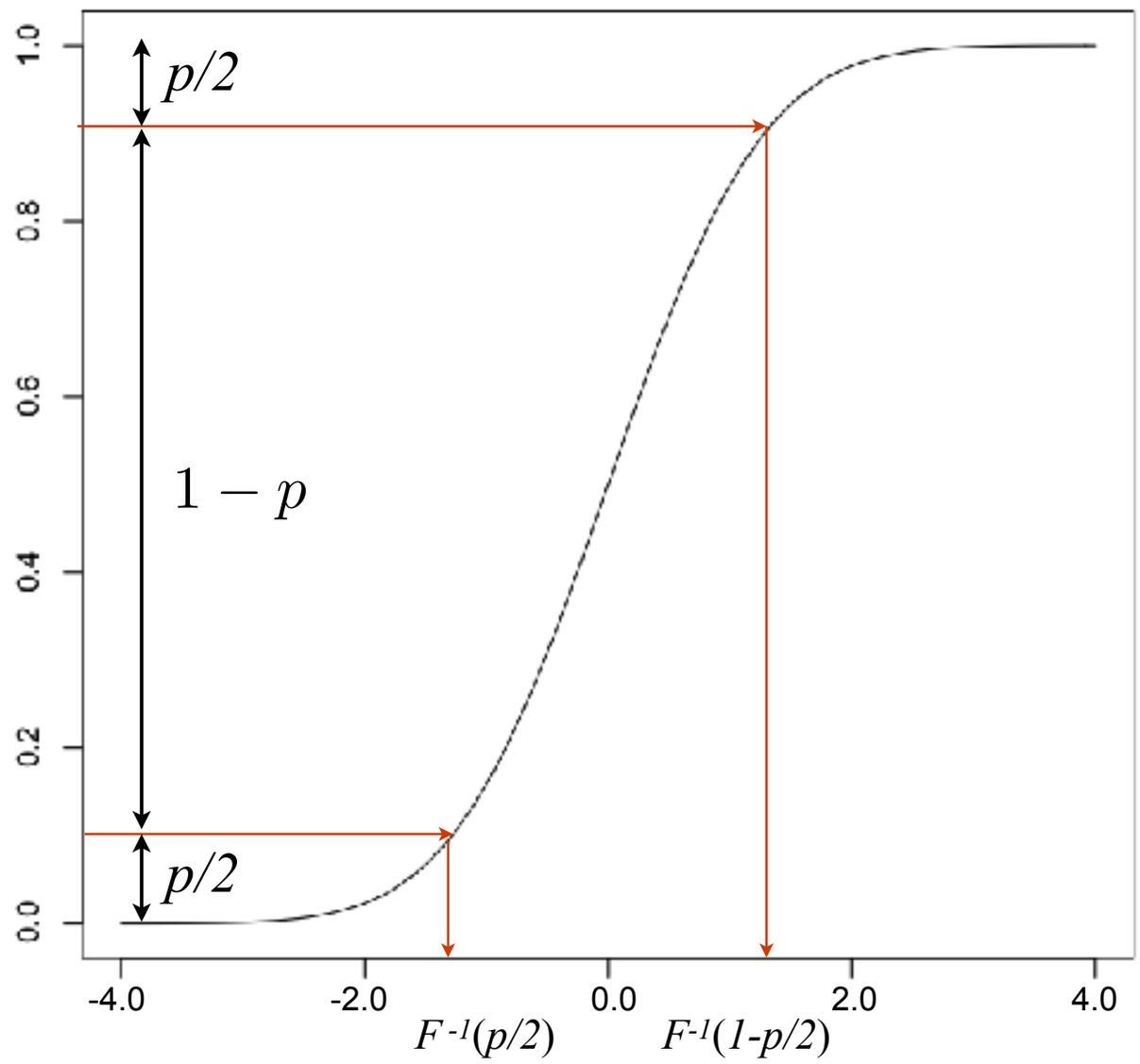
Jaký je vztah mezi $x_{p/2}$ a $y_{p/2}$?

... označme tedy

$$\Phi^{-1}(1-p/2) = u_p$$

... potom

$$\begin{aligned} \Phi^{-1}(p/2) &= -u_p \\ F^{-1}(1-p/2) &= \mu + \sigma u_p \\ F^{-1}(p/2) &= \mu - \sigma u_p \end{aligned}$$



Tedy pro $\alpha \in (0,1)$

$$X \sim N(0,1) \Rightarrow \qquad P(-u_\alpha \leq X \leq u_\alpha) = 1-\alpha$$

$$Y \sim N(\mu, \sigma^2) \Rightarrow \qquad P(\mu - \sigma u_\alpha \leq Y \leq \mu + \sigma u_\alpha) = 1-\alpha$$

$$P\left(-u_\alpha \leq \frac{Y-\mu}{\sigma} \leq u_\alpha\right) = 1-\alpha$$

Normální (Gaussovo) rozdělení

Střední hodnota náhodné veličiny X:

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

Rozptyl náhodné veličiny X:

$$var(X) = E(X - E(X))^2 = E(X^2) - \big(E(X)\big)^2$$

$$X \sim N(\mu, \sigma^2) \Rightarrow \quad E(X) = \mu \qquad var(X) = \sigma^2$$

Jak tyto parametry najít, když máme k dispozici pouze několik měření veličiny X?

$$x_1, x_2, x_3, \ldots, x_n \qquad\qquad X_1, X_2, X_3, \ldots, X_n$$