

Plánování Průmyslových Experimentů

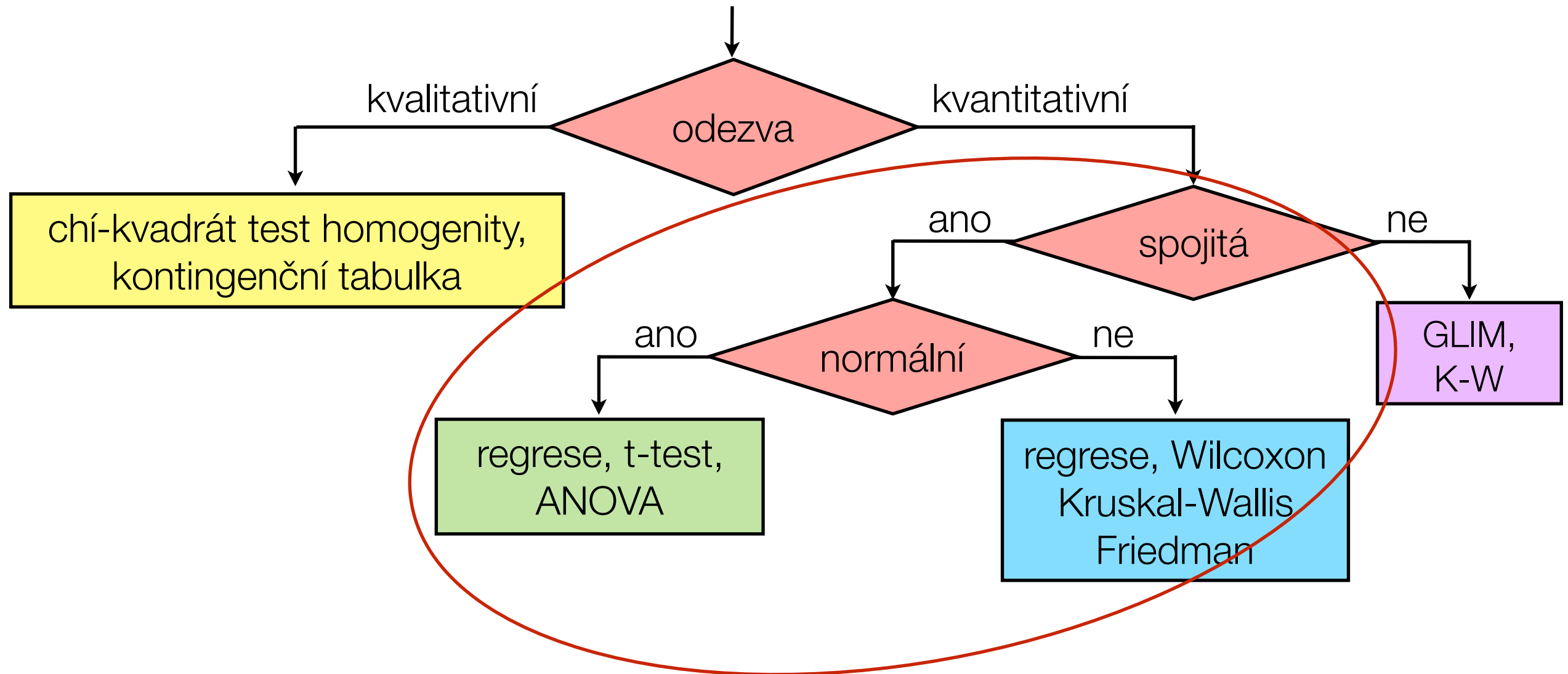
Část 2. Metody vyhodnocování experimentů

- Vyhodnocení experimentu: stanovení charakteristik procesu a jeho optimalizace, vyhodnocení vlastností materiálů zpravidla metodami matematické statistiky

Klíčové pojmy: Testování hypotéz, odhady parametrů, regrese, korelace



2.2. Kvantitativní odezva



2.2. Vsuvka 2: Normální rozdělení

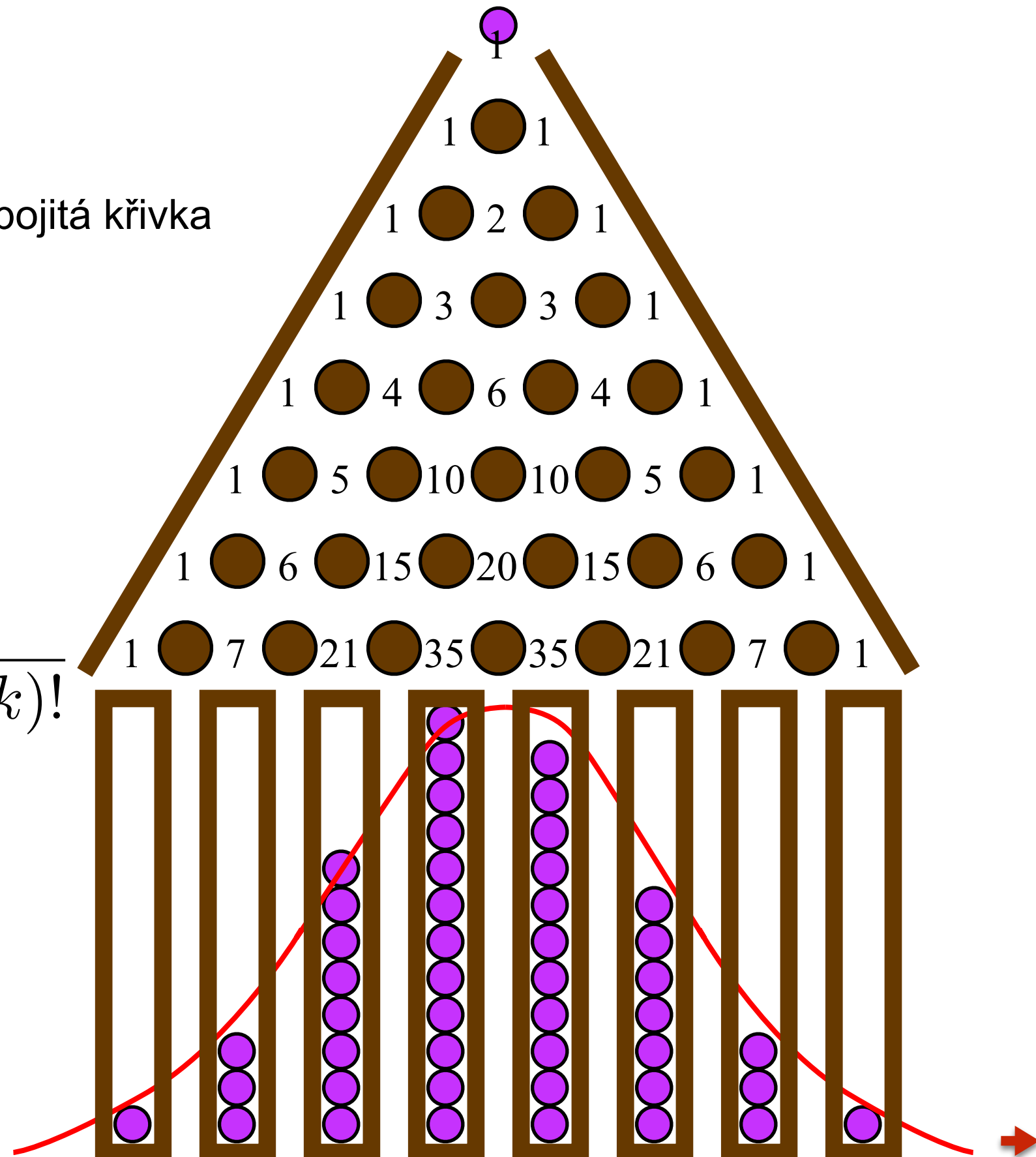
1733 : Abraham de Moivre

pokusy s mincemi $\longrightarrow$ spojitá křivka

Galtonova deska

Pascalův trojúhelník

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



Moivre navrhl rovnici:

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

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

2.2. Vsuvka 2: Normální rozdělení

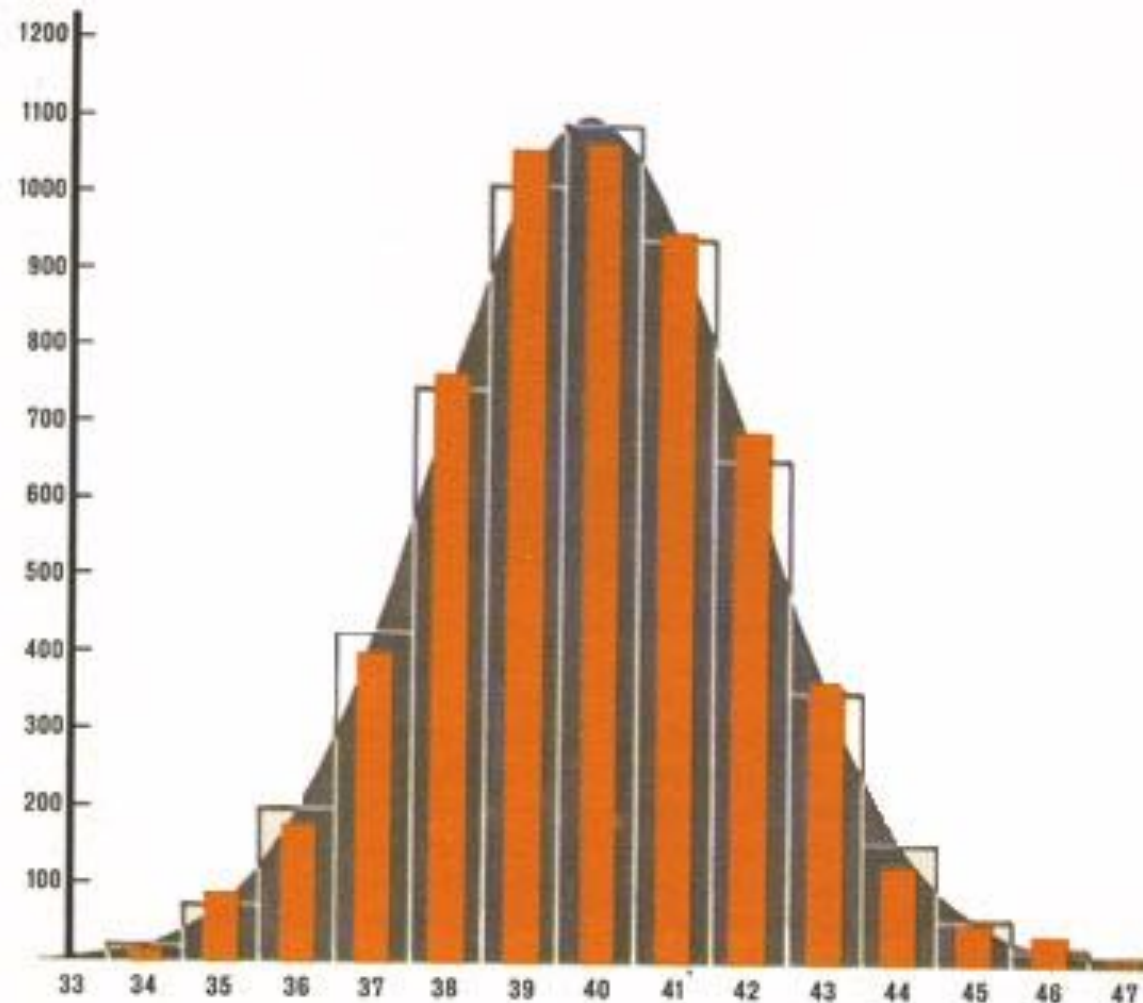
přelom 18. a 19. století: rozdíly v astronomických měřeních
- která hodnota je správná?

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

gaussova křivka

počátek 19. století: Adolphe Quételet

- antropometrická měření anglických vojáků



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

2.2. Vsuvka 2: Normální rozdělení

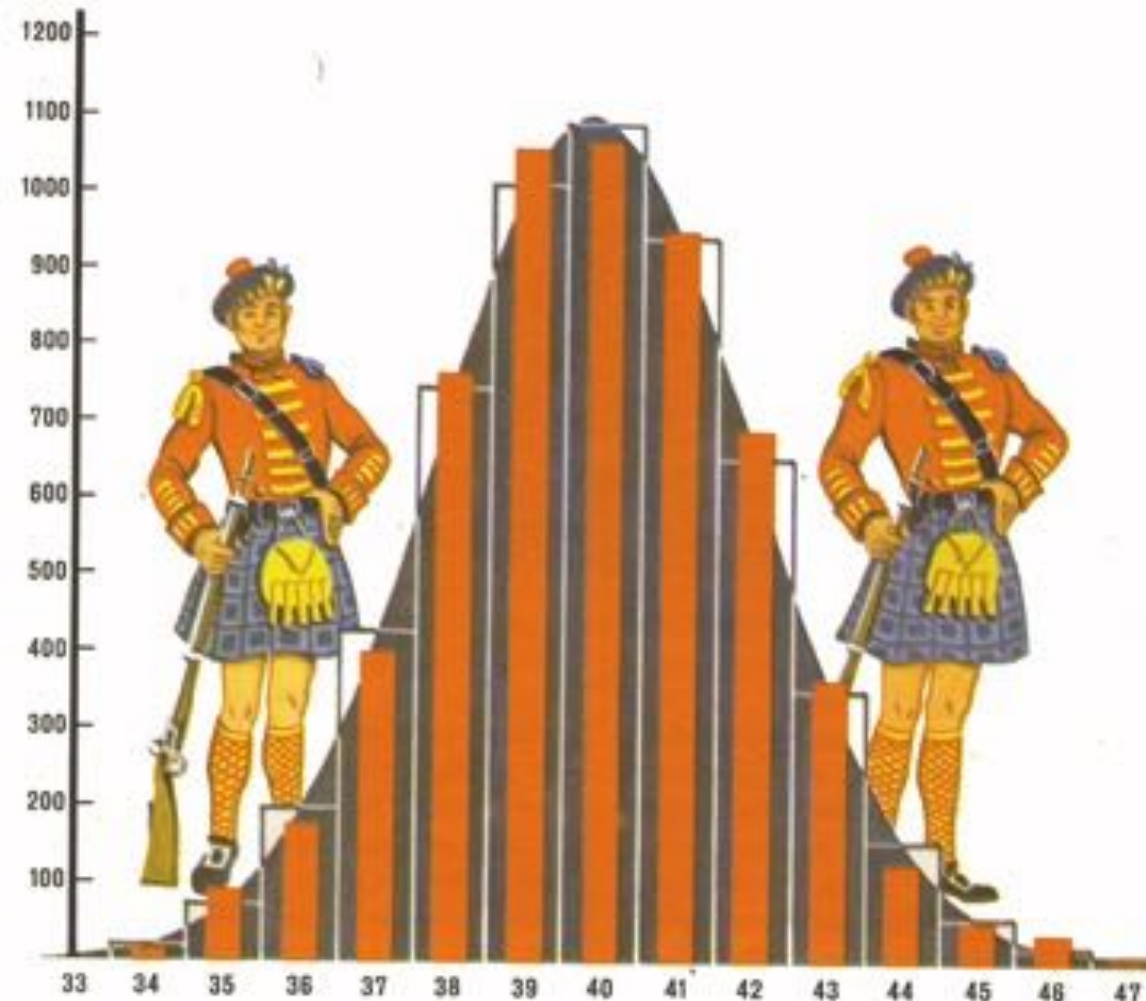
přelom 18. a 19. století: rozdíly v astronomických měřeních
- která hodnota je správná?

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

gaussova křivka

počátek 19. století: Adolphe Quételet

- antropometrická měření anglických vojáků



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

normální rozdělení

2.2. Vsuška 2: Normální rozdění

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

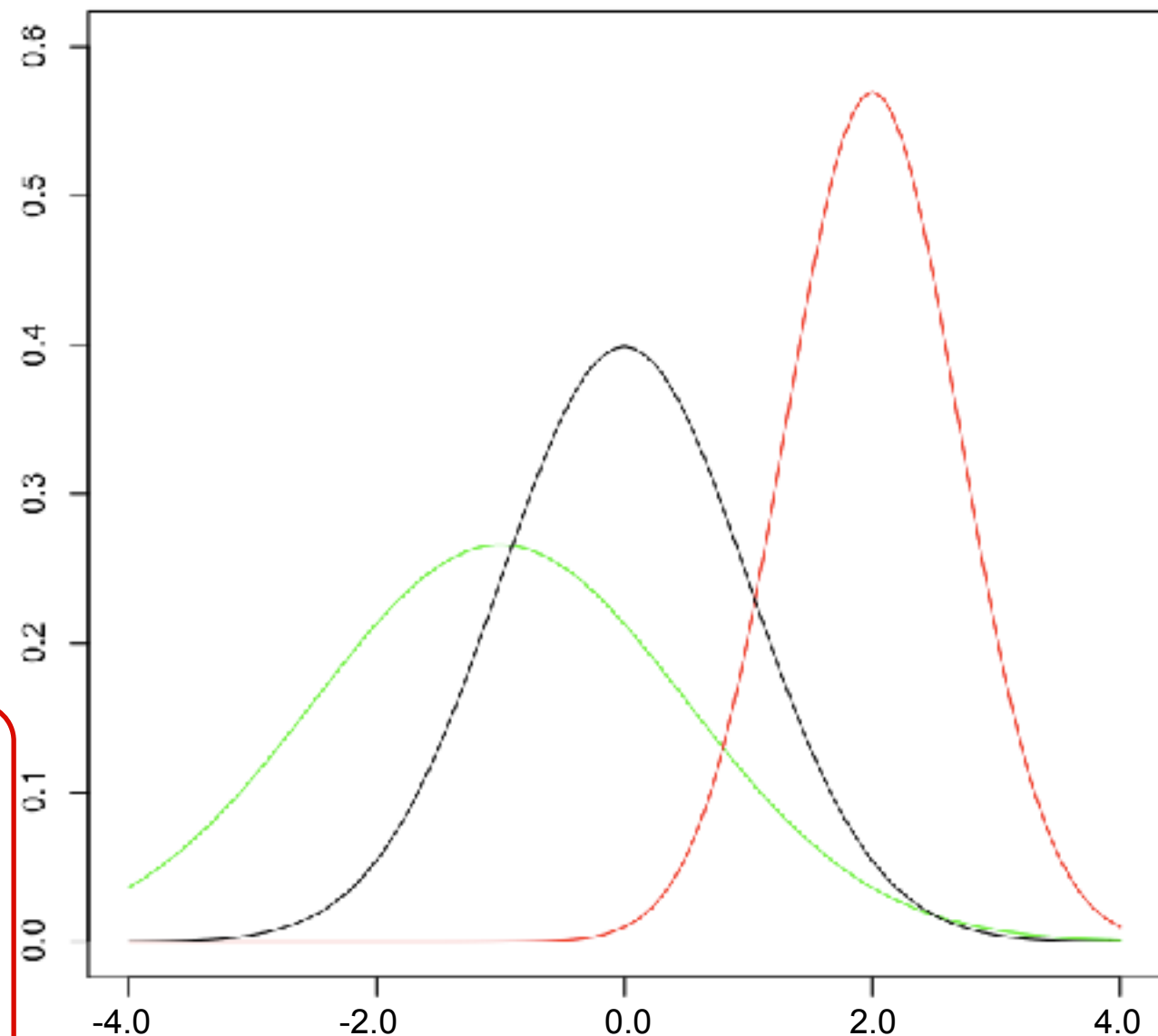
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

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

hustota pravděpodobnosti

$\sigma > 0$ - parametr měřítka,

$\mu \in R$ - parametr polohy



Normální rozdění $N(\mu, \sigma^2)$

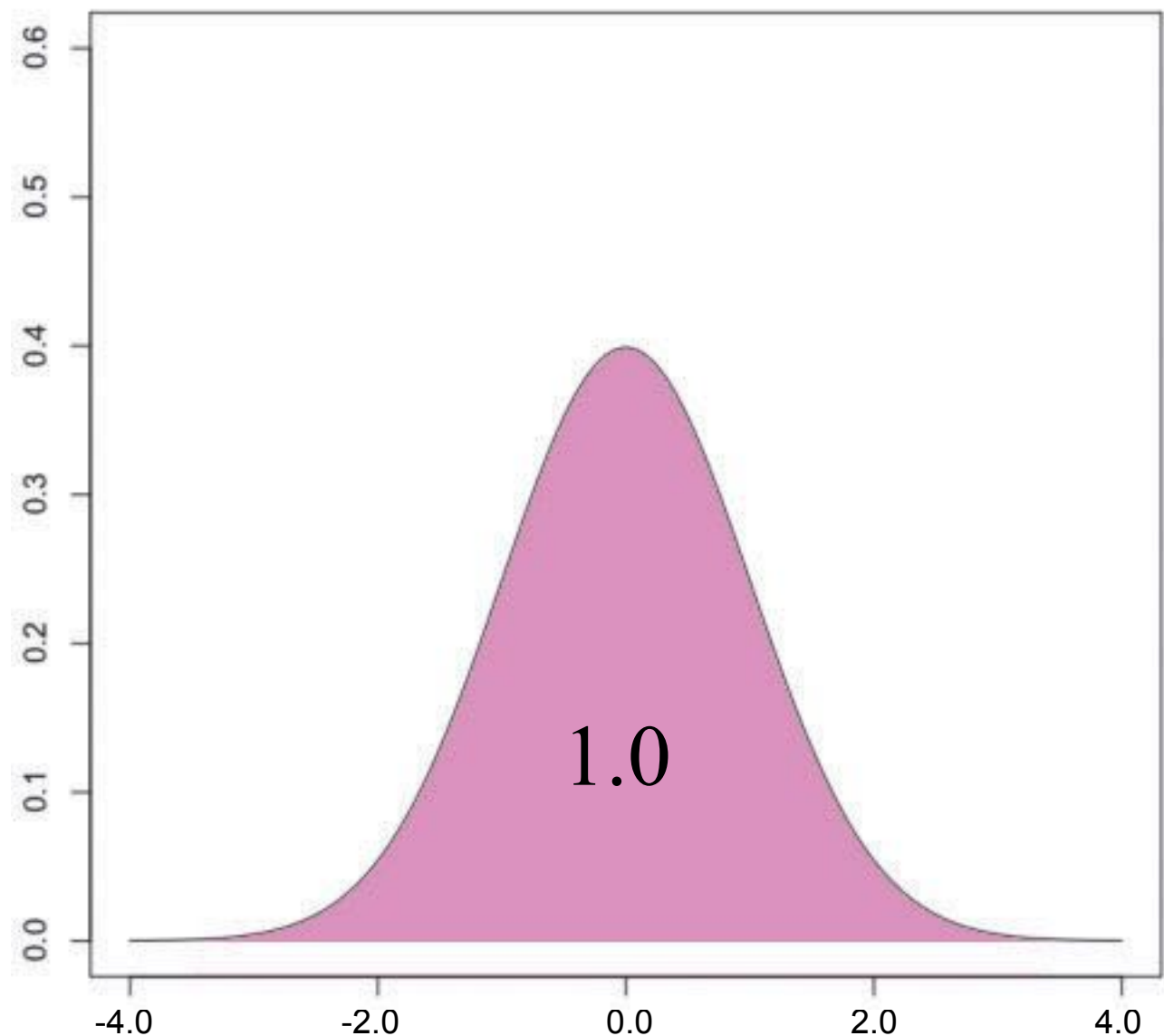
2.2. Vsuvka 2: Normální rozdělení

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

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

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

Standardní normální
rozdělení



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

2.2. Vstupka 2: Normální rozdělení

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

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

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

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

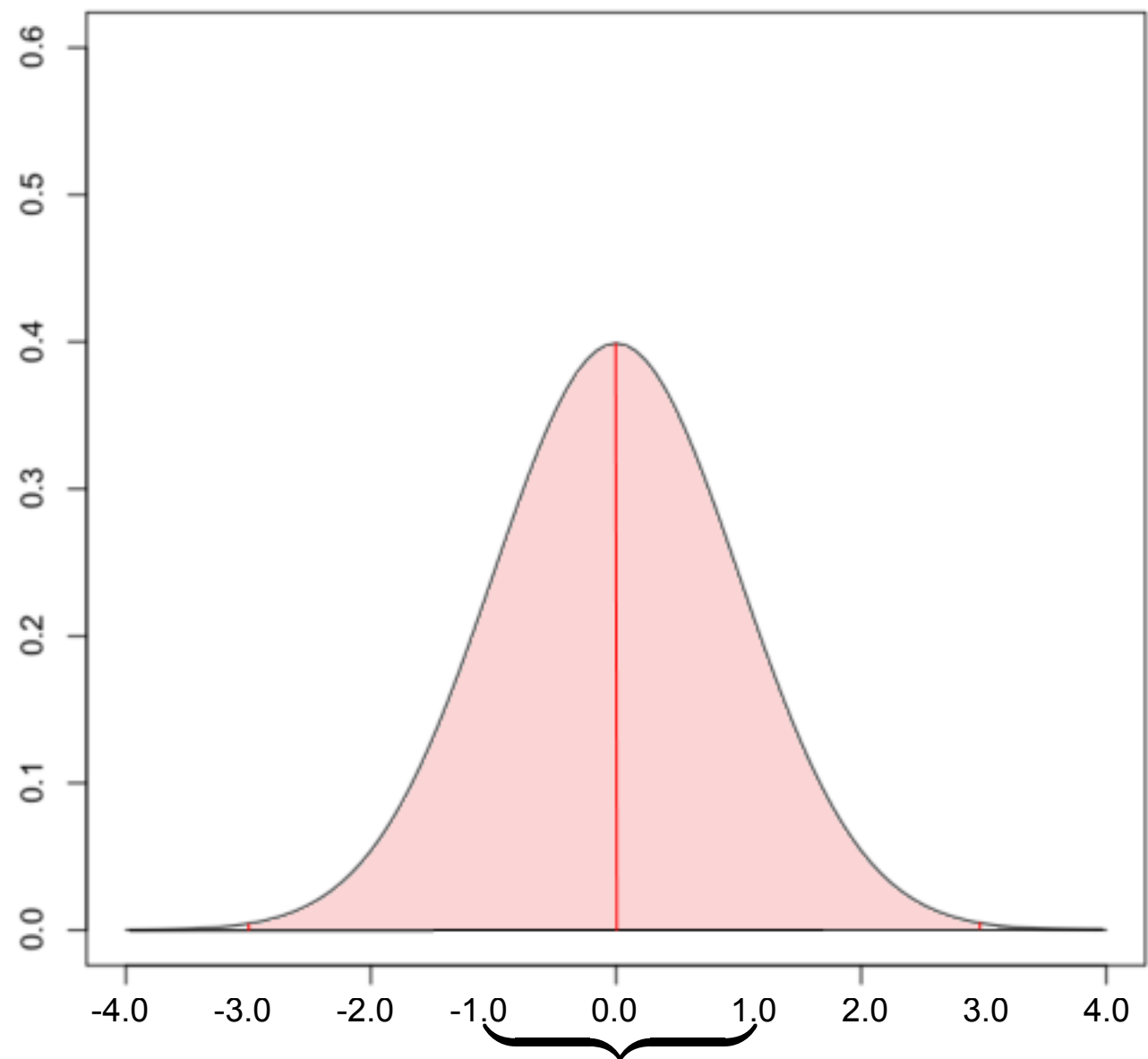
$$P(-3 \leq X \leq 3) = 0.9973$$

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

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

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

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



68,26%

95%

99.7%

2.2. Vsuvka 2: Normální rozdělení

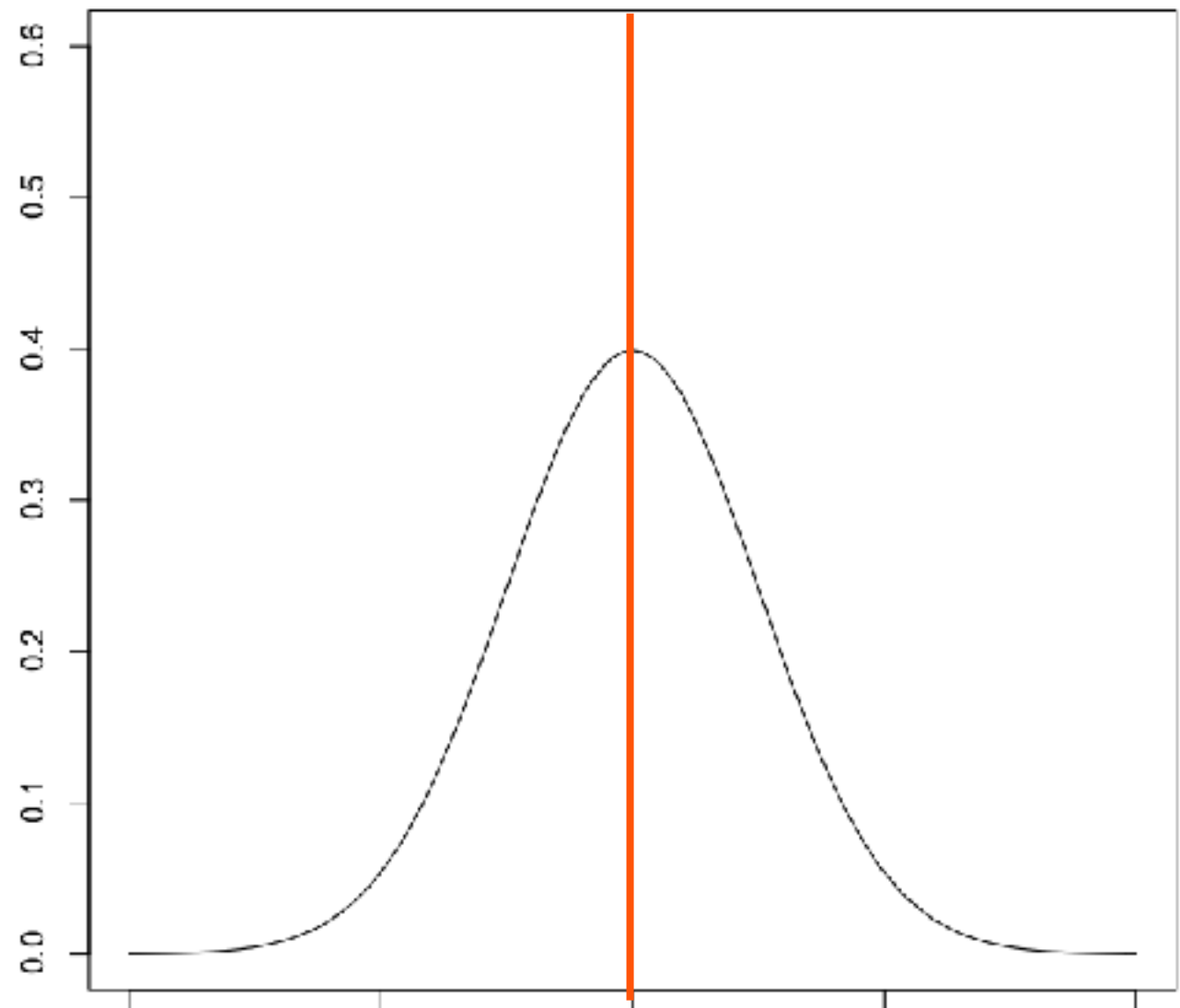
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$
$$x \in \mathbb{R}$$

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

Symetrie: $f(x) = f(-x)$

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

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



2.2. Vsuvka 2: Normální rozdělení

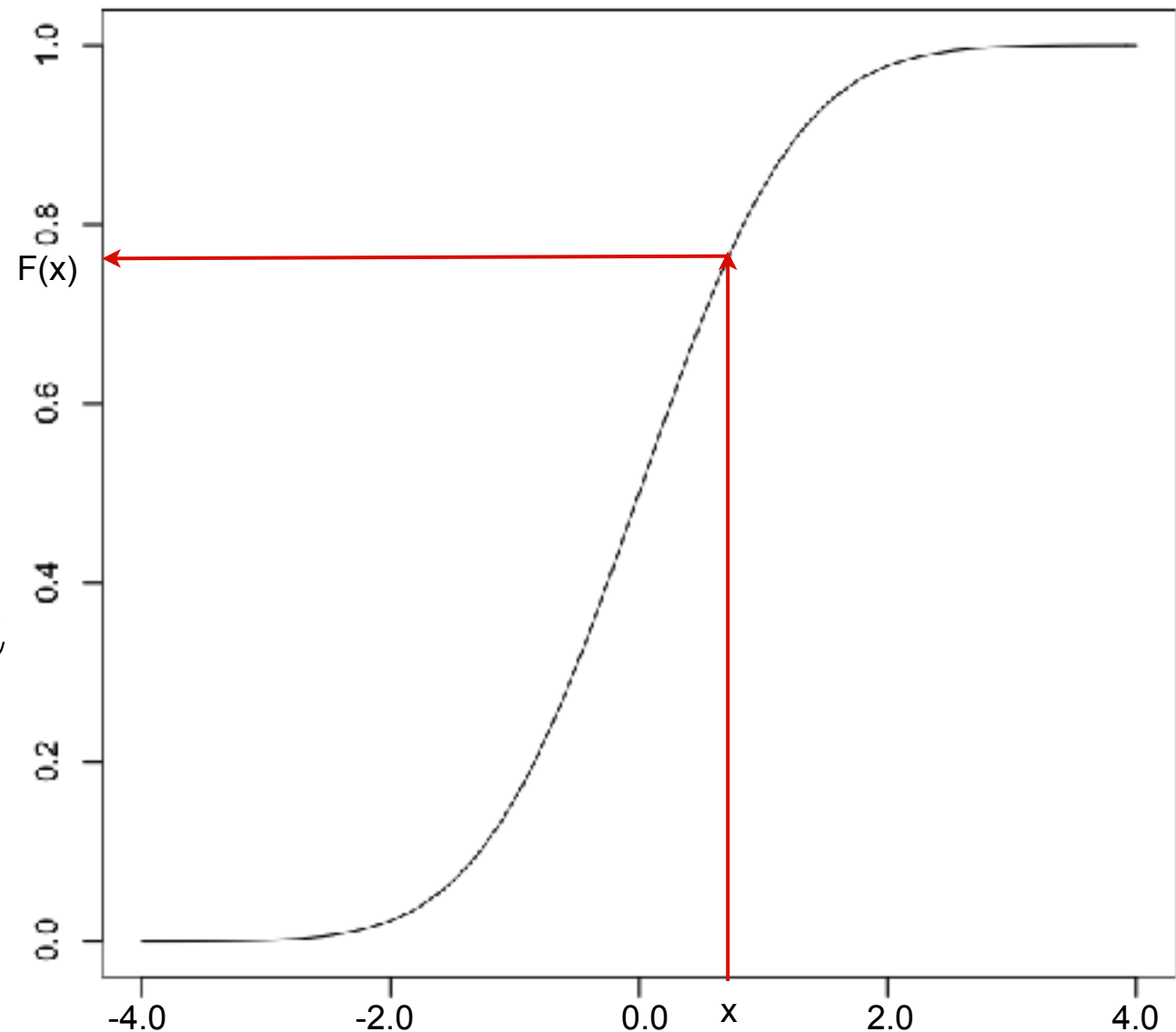
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$
$$x \in \mathbb{R}$$

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

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

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

distribuční funkce



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 ?

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

2.2. Vstupka 2: Normální rozdělení

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

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

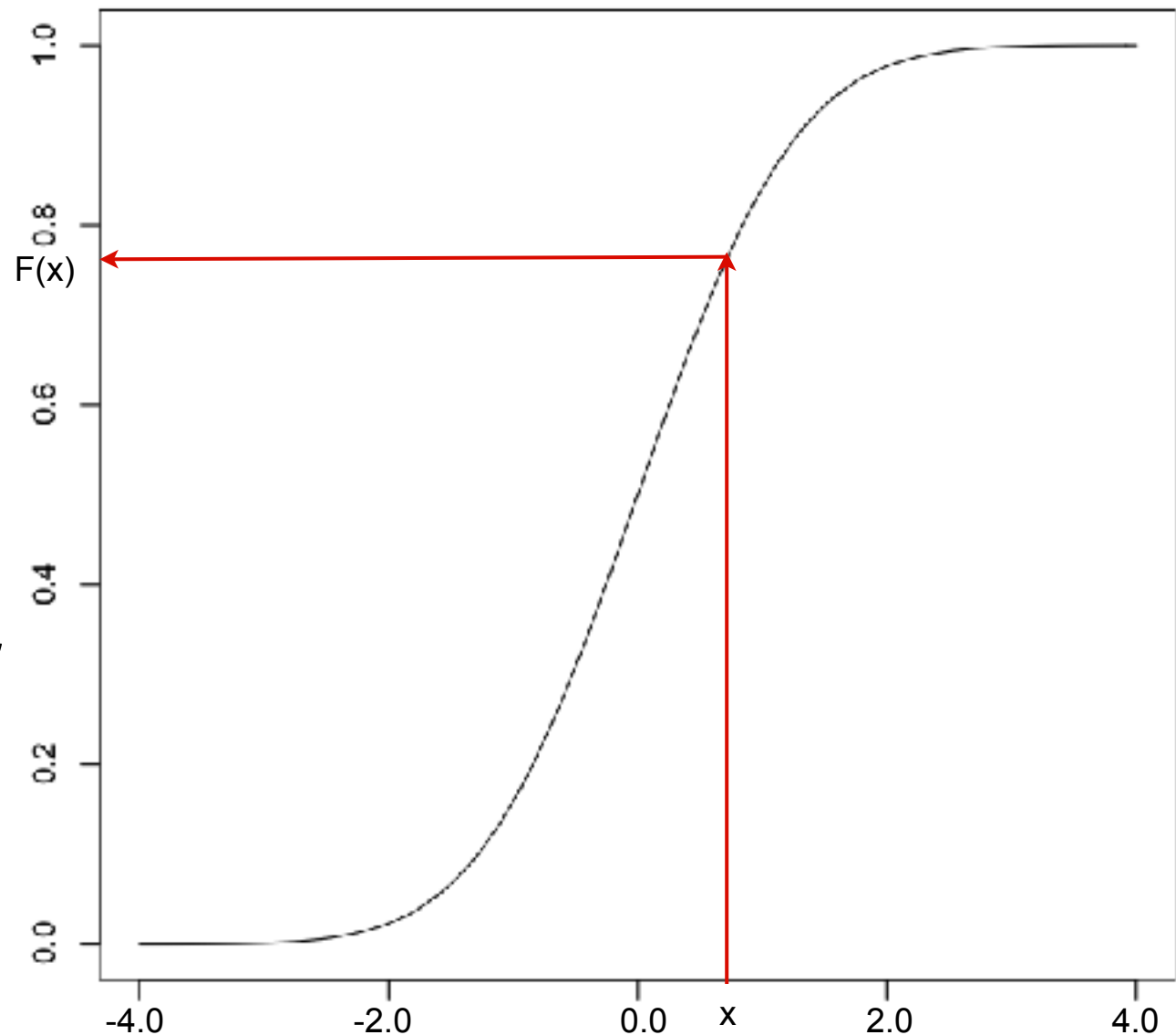
$$P(X \leq x) = \int_{-\infty}^x f(u) du$$

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

distribuční funkce

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

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



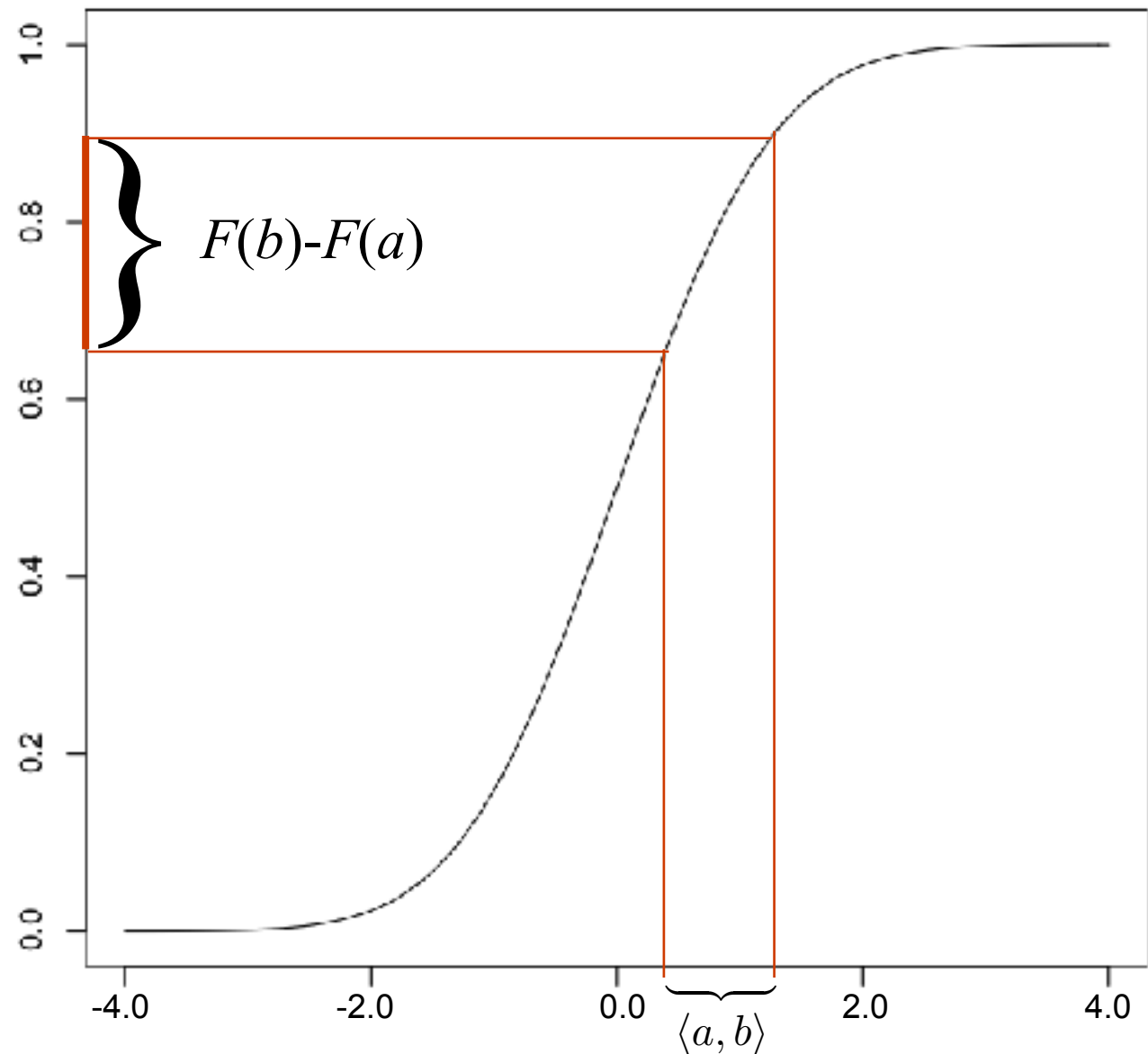
2.2. Vsuvka 2: Normální rozdělení

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

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

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

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



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

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

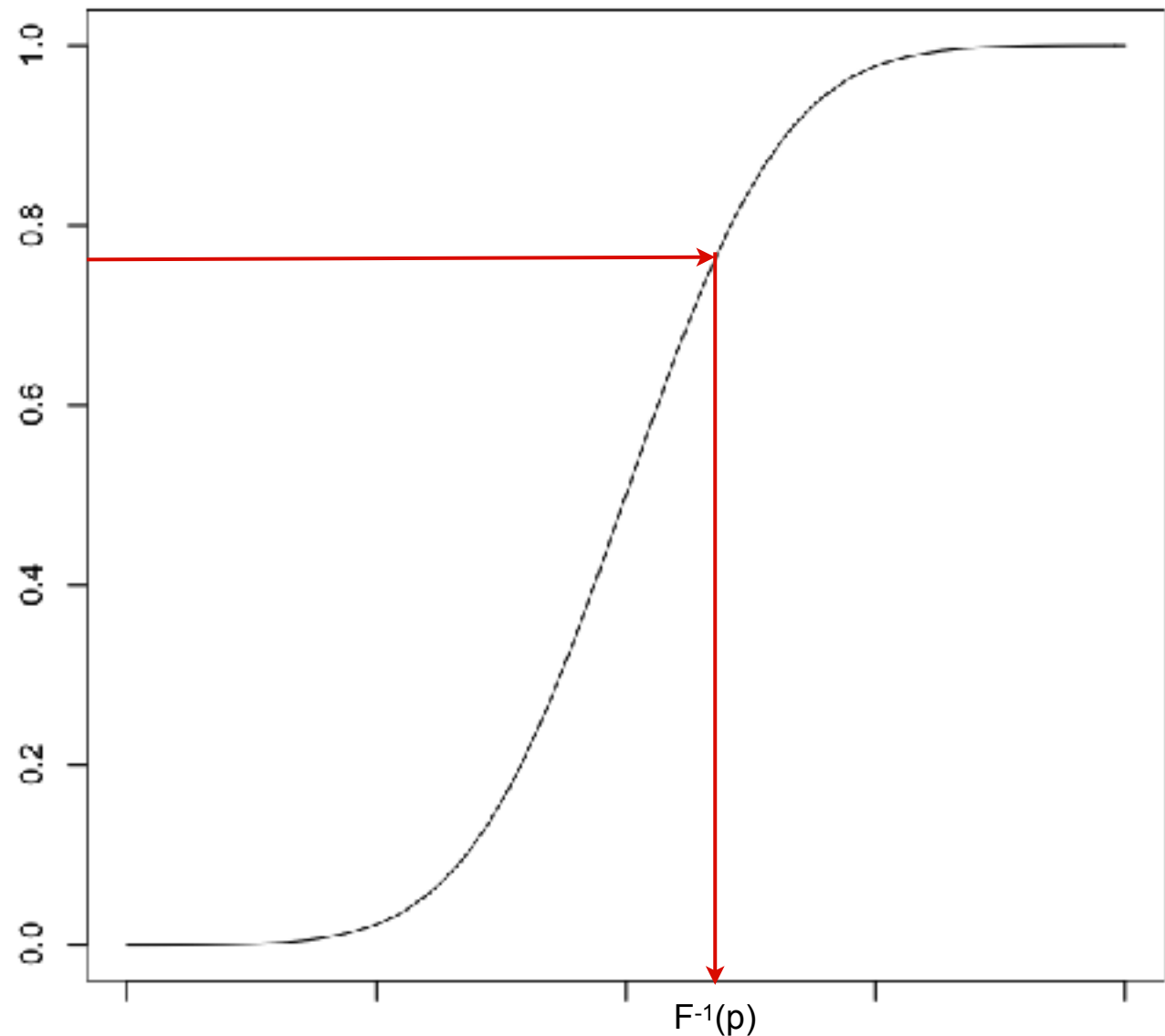
2.2. Vsuvka 2: Normální rozdělení

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

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

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

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



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 ?

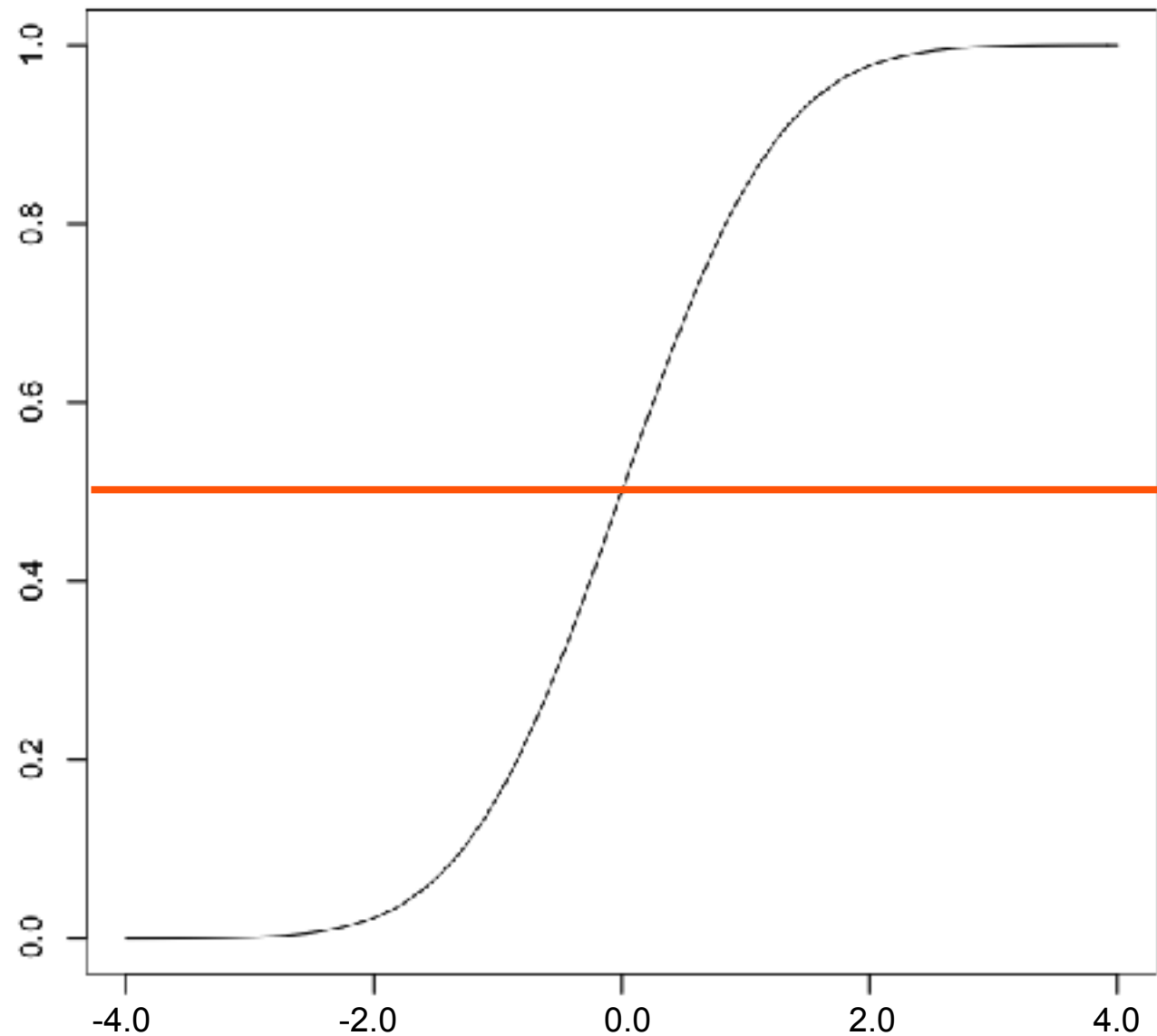
2.2. Vstupka 2: Normální rozdělení

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

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

Symetrie distribuční funkce:

$$\Phi(x) = 1 - \Phi(-x)$$



2.2. Vstupka 2: Normální rozdělení

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

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

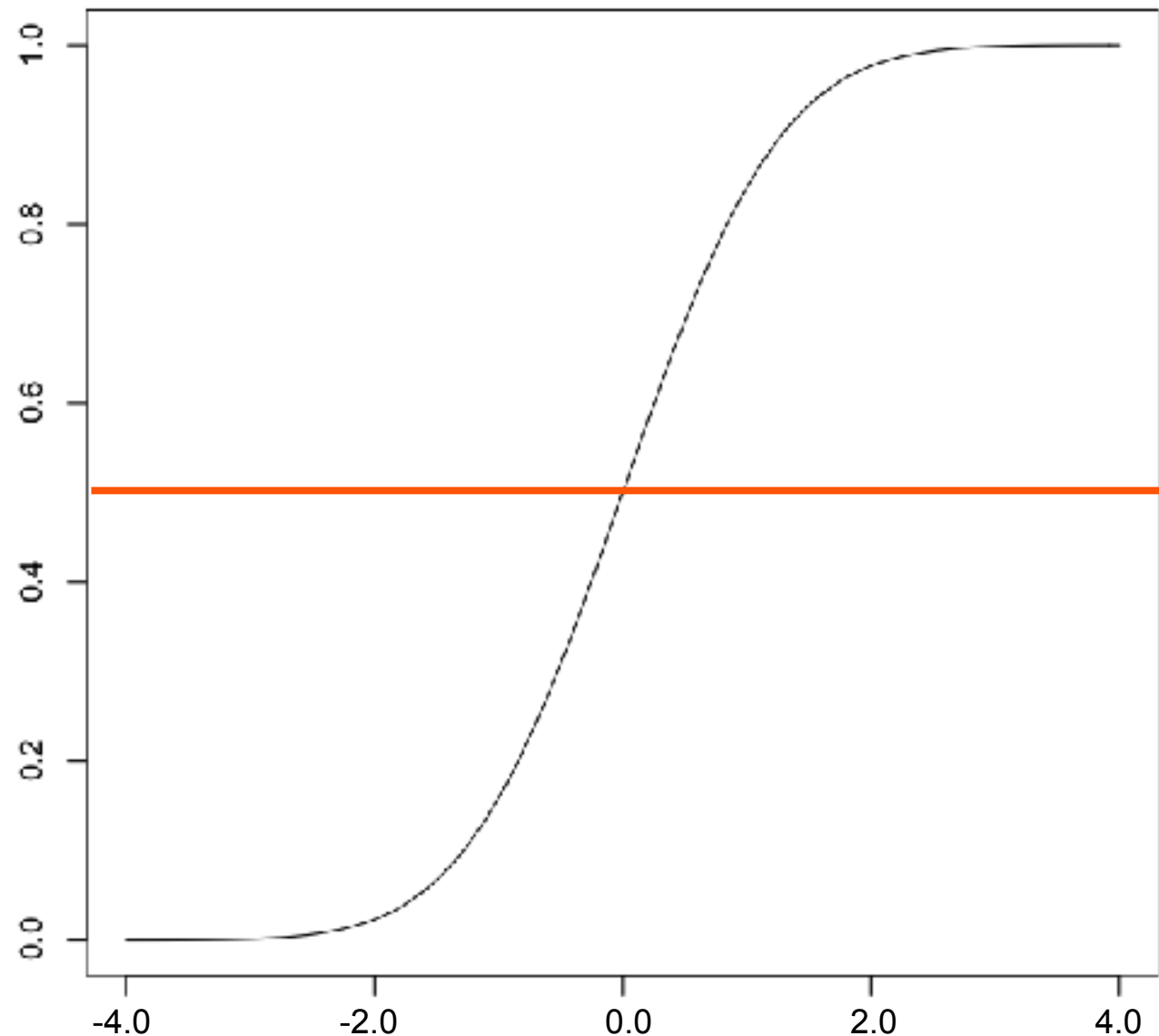
Symetrie distribuční funkce:

$$\Phi(x) = 1 - \Phi(-x)$$

Symetrie kvantilové funkce:

$u_{p/2} = \Phi^{-1}(p/2)$ - kvantil standardního normálního rozdělení a

$u_{1-p/2} = \Phi^{-1}(1-p/2)$ - kvantil standardního normálního rozdělení



$$u_{p/2} = -u_{1-p/2}$$

2.2. Vsuvka 2: Normální rozdělení

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

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

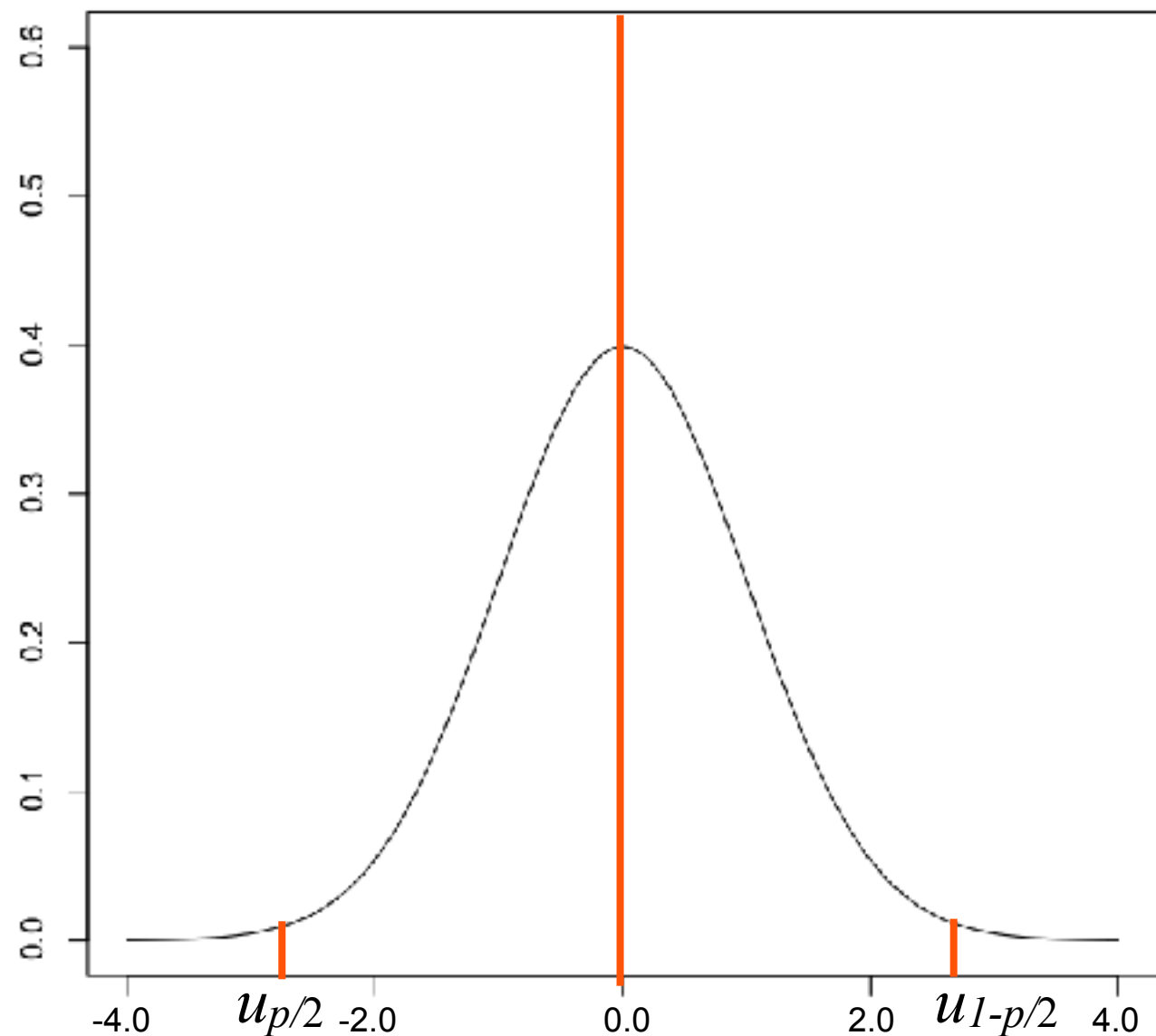
Symetrie distribuční funkce:

$$\Phi(x) = 1 - \Phi(-x)$$

Symetrie kvantilové funkce:

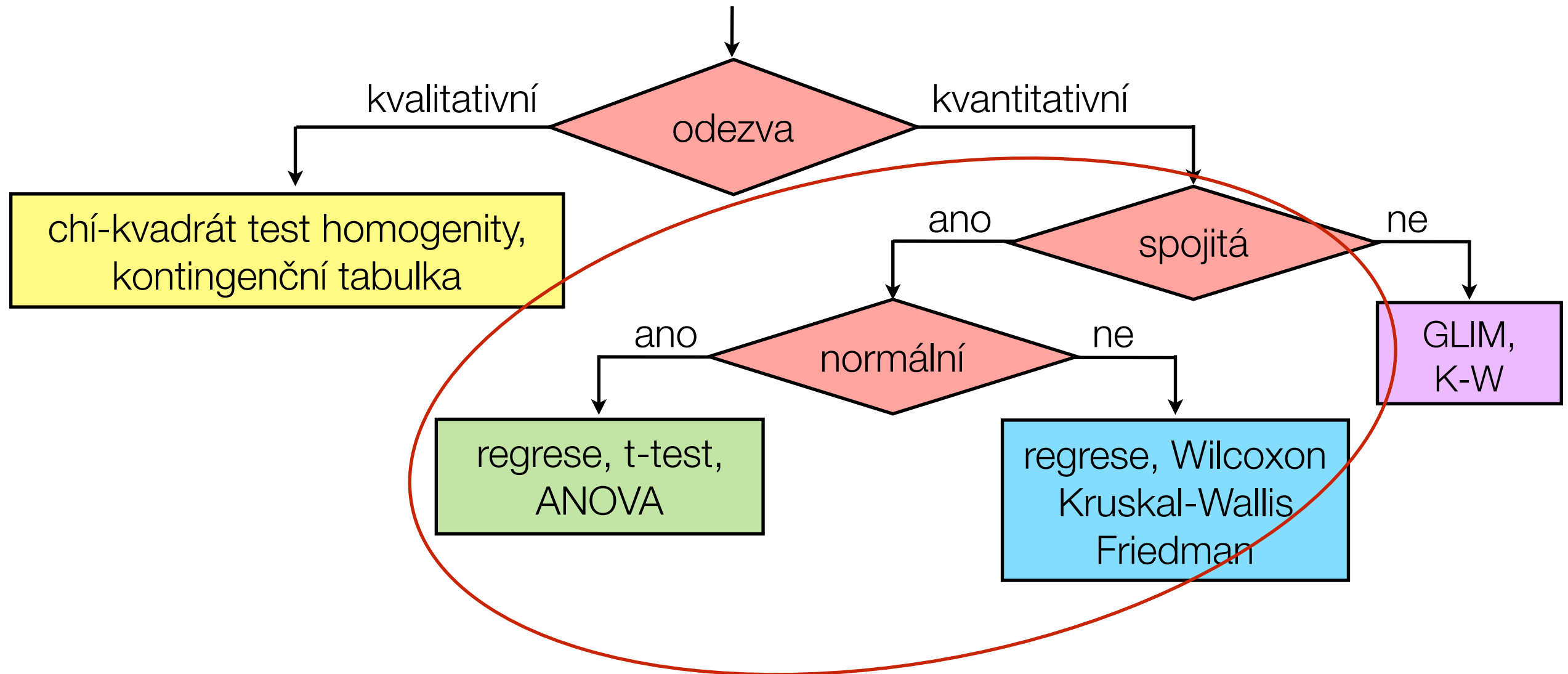
$u_{p/2} = \Phi^{-1}(p/2)$ - kvantil standardního normálního rozdělení a

$u_{1-p/2} = \Phi^{-1}(1-p/2)$ - kvantil standardního normálního rozdělení



$$u_{p/2} = -u_{1-p/2}$$

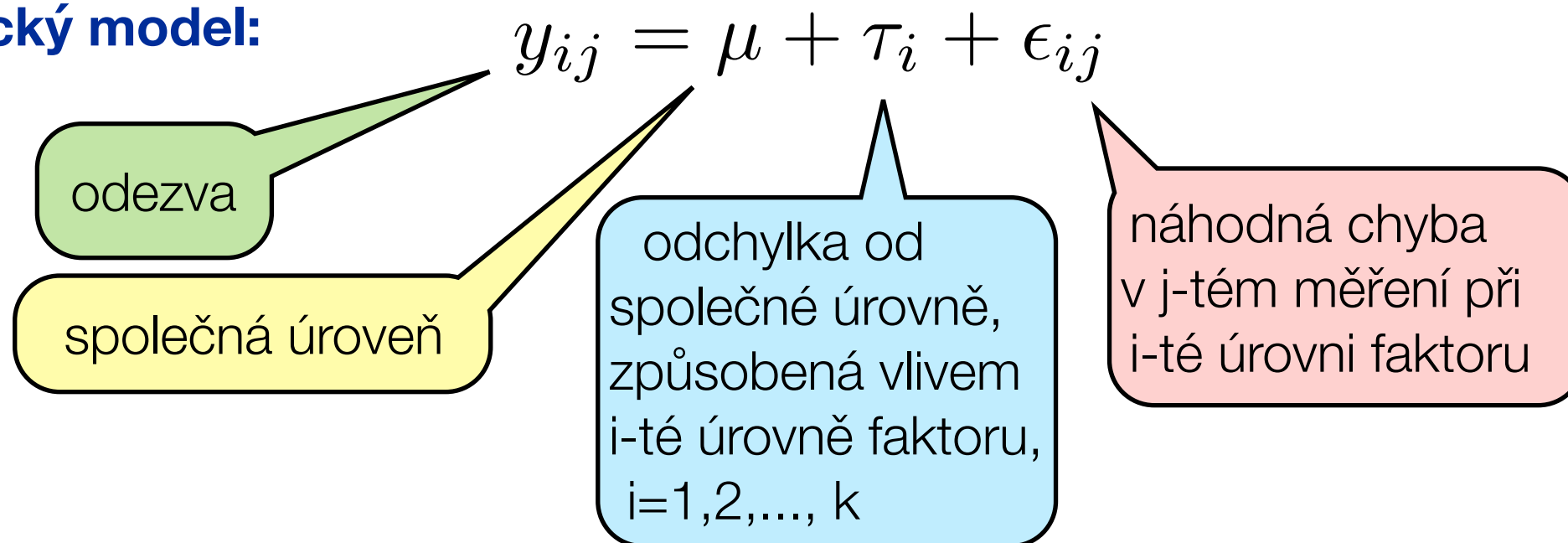
2.2. Kvantitativní odezva



2.2. Kvantitativní odezva, jeden faktor

- uvažujeme vliv pouze jediného (hlavního) faktoru X
- tento faktor může mít k úrovní
- pro každou úroveň provedeme n_i měření (replikací), $i=1,2,\dots,k$
- měření můžeme sdružovat do bloků podle nějakého vedlejšího (blokového) faktoru
- provádíme znáhodnění pořadí měření (buď přes celý experiment nebo pouze uvnitř bloků)

Statistický model:



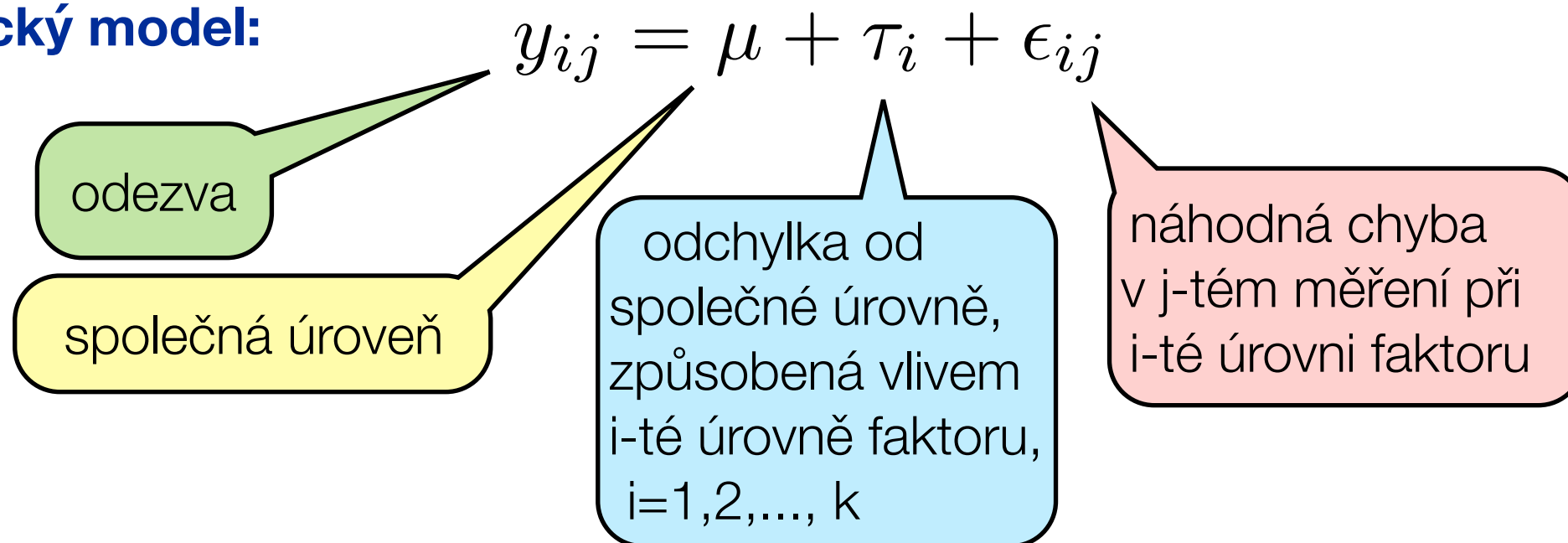
**O náhodných chybách ϵ_{ij} předpokládáme, že jsou nezávislé, stejně rozdělené s nulovou střední hodnotou a stejným rozptylem σ^2 .
Speciálně, předpokládáme rozdělení $N(0, \sigma^2)$**



2.2. Kvantitativní odezva, jeden faktor

- uvažujeme vliv pouze jediného (hlavního) faktoru X
- tento faktor může mít k úrovní
- pro každou úroveň provedeme n_i měření (replikací), $i=1,2,\dots,k$
- měření můžeme sdružovat do bloků podle nějakého vedlejšího (blokového) faktoru
- provádíme znáhodnění pořadí měření (buď přes celý experiment nebo pouze uvnitř bloků)

Statistický model:



vliv i-té úrovně faktoru, $i=1,2,\dots, k$: $\mu_i = \mu + \tau_i$



2.2. Kvantitativní odezva, jeden faktor

pořadové číslo měření (replikace)

úroveň faktoru (ošetření)

	1	2	3	...	n
1	y_{11}	y_{12}	y_{13}	...	y_{1n_1}
2	y_{21}	y_{22}	y_{23}	...	y_{2n_2}
3	y_{31}	y_{32}	y_{33}	...	y_{3n_3}
:	:	:	:	$\ddots$	:
k	y_{k1}	y_{k2}	y_{k3}	...	y_{kn_k}



2.2. Kvantitativní odezva, jeden faktor

Jeden faktor o dvou úrovních: Srovnání dvou souborů dat, normální rozdělení

1. úroveň faktoru, n replikací: $X_1, X_2, \dots, X_n$ $X \sim N(\mu_X, \sigma_X^2)$
2. úroveň faktoru, m replikací: $Y_1, Y_2, \dots, Y_m$ $Y \sim N(\mu_Y, \sigma_Y^2)$

	1	2	3	...	n
A	x_1	x_2	x_3	...	x_n
B	y_1	y_2	y_3	...	y_m

- a) dvě nezávislá měření
- b) párová měření



2.2. Vsuvka 3: Analýza datového souboru

- Základní charakteristiky
 - pořadové statistiky: minimum, maximum, medián, kvartily
 - momenty: průměr, výběrový rozptyl, šikmost, špičatost
 - modus, rozpětí
- Grafická analýza
 - krabicový graf
 - histogram
 - Q-Q graf
 - rozptylový graf
- Test normality



2.2. Vsuvka 3: Analýza datového souboru

- Základní charakteristiky
 - pořadové statistiky: minimum, maximum, medián, kvartily
 - momenty: průměr, výběrový rozptyl, šikmost, špičatost
 - modus, rozpětí

1. výběrový moment = aritmetický průměr:
(bodový odhad střední hodnoty) $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
2. výběrový centrální moment = výběrový rozptyl $m_2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$
3. výběrový centrální moment $m_3 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^3$
(bodový odhad
koeficientu šikmosti: $S_{kew} = \frac{m_3}{\sqrt{m_2^3}}$
4. výběrový centrální moment $m_4 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^4$
(bodový odhad
koeficientu špičatosti: $K_{urt} = \frac{m_4}{m_2^2} - 3$



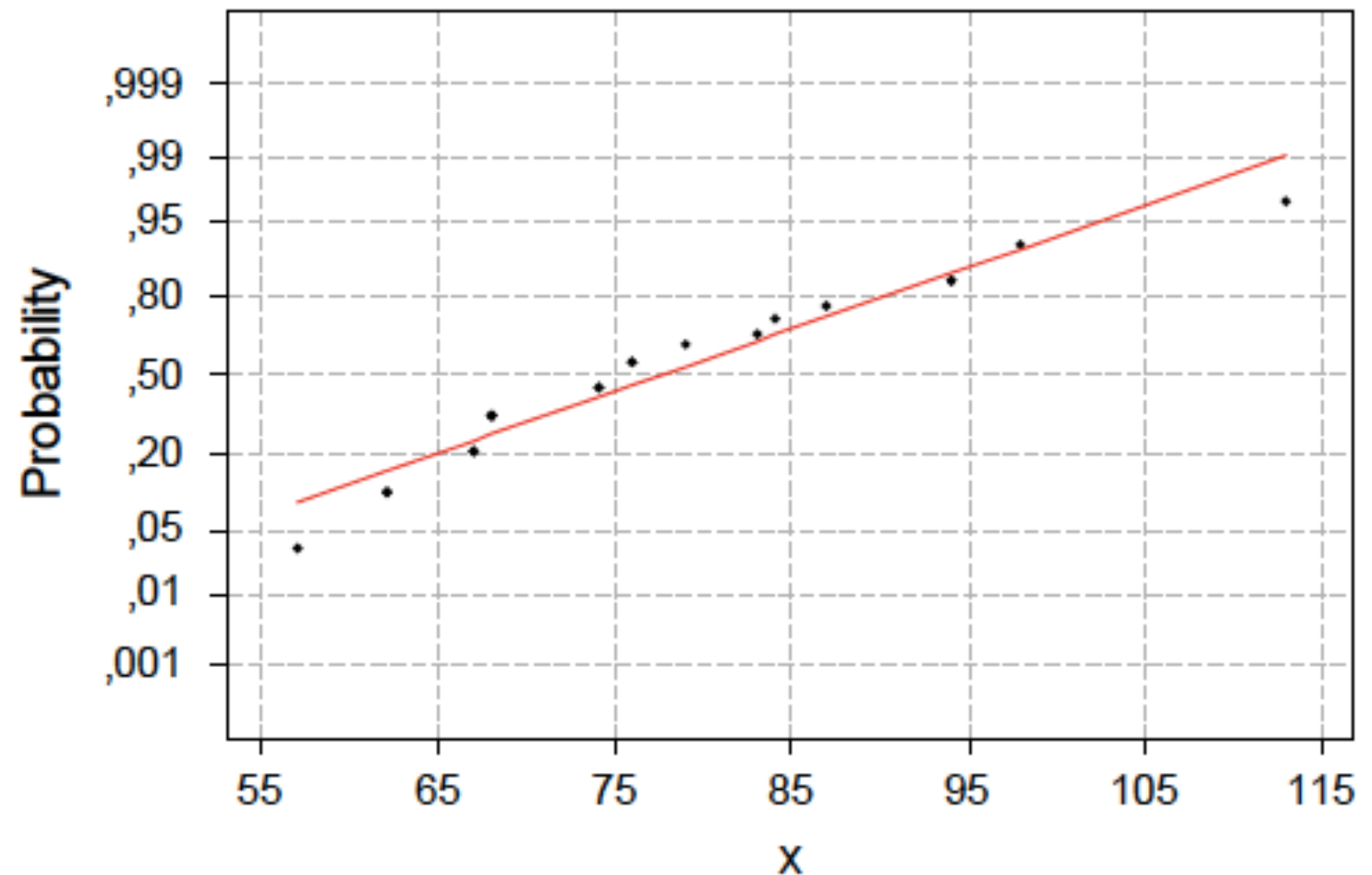
2.2. Vsuvka 3: Analýza datového souboru



2.2. Vsuvka 3: Analýza datového souboru

- Základní charakteristiky
 - pořadové statistiky: minimum, maximum, medián, kvartily
 - momenty: průměr
 - modus, rozpětí
- **Grafická analýza**
 - krabicový graf
 - histogram
 - Q-Q graf

Normal Probability Plot

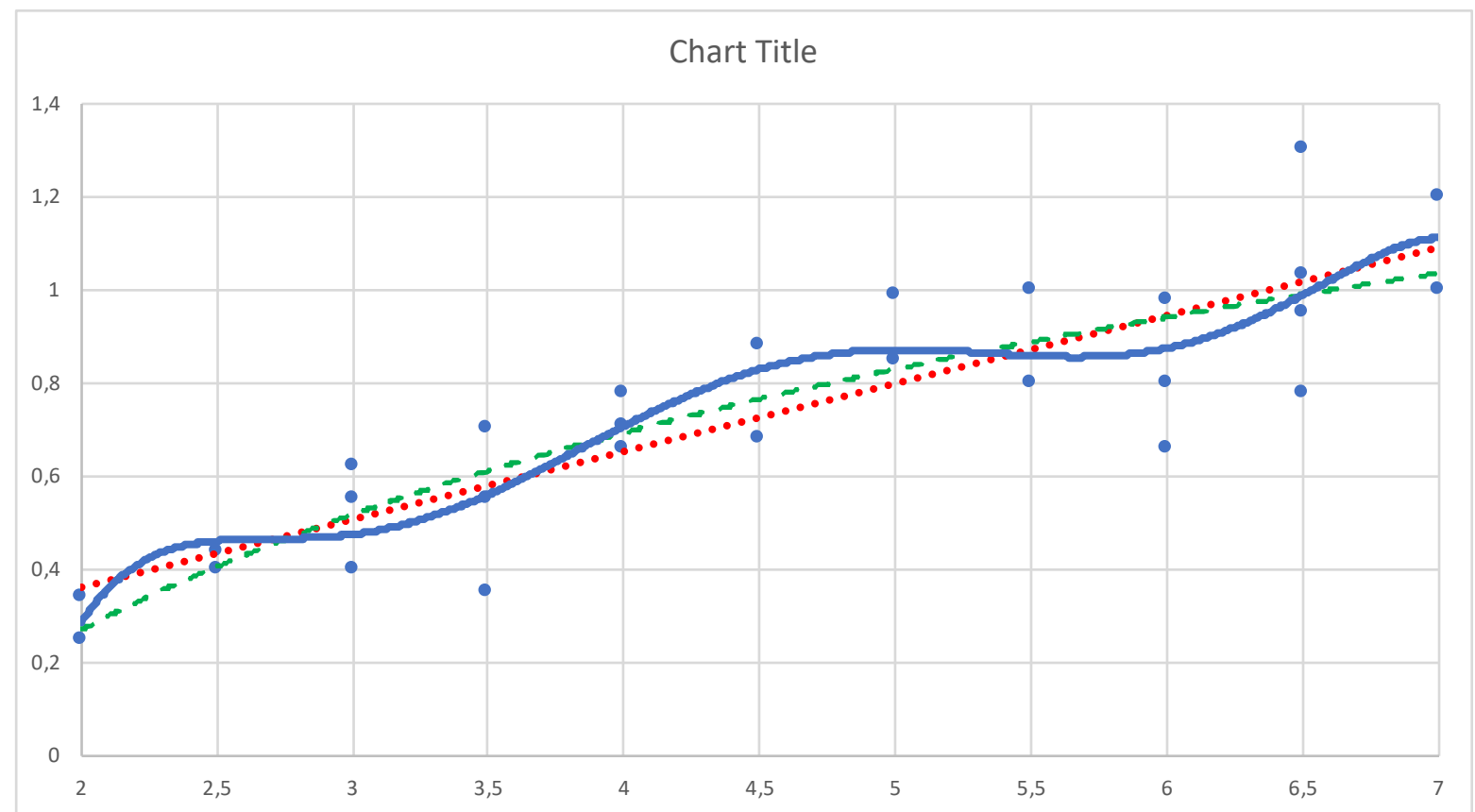


Average: 77,55
StDev: 14,1625
N: 20

Anderson-Darling Normality Test
A-Squared: 0,455
P-Value: 0,240

2.2. Vsuvka 3: Analýza datového souboru

- Základní charakteristiky
 - pořadové statistiky: minimum, maximum, medián, kvartily
 - momenty: průměr, výběrový rozptyl, šikmost, špičatost
 - modus, rozpětí
- **Grafická analýza**
 - krabicový graf
 - histogram
 - Q-Q graf
 - rozptylový graf



2.2. Vsuvka 3: Analýza datového souboru

- Základní charakteristiky
 - pořadové statistiky: minimum, maximum, medián, kvartily
 - momenty: průměr, výběrový rozptyl, šikmost, špičatost
 - modus, rozpětí
- Grafická analýza
 - krabicový graf
 - histogram
 - Q-Q graf
 - rozptylový graf
- **Test normality**



2.2. Vsuvka 3: Analýza datového souboru

Test normality na základě šikmosti a špičatosti

Za předpokladu, že výběr pochází z normálního rozdělení,
platí pro index šikmosti:

$$E(S_{kew}^{norm}) = 0$$

$$Var(S_{kew}^{norm}) = \frac{6(n-2)}{(n+1)(n+3)}$$

a pro index špičatosti:

$$E(K_{urt}^{norm}) = 3 - \frac{6}{n+1}$$

$$Var(K_{urt}^{norm}) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$$

Máme-li dostatečný počet pozorování (řádově stovky), mají statistiky

$$T_3 = \frac{S_{kew}^{norm}}{\sqrt{Var(S_{kew}^{norm})}} \quad T_4 = \frac{K_{urt}^{norm} - E(K_{urt}^{norm})}{\sqrt{Var(K_{urt}^{norm})}}$$

přibližně standardní normální rozdělení pravděpodobnosti.



2.2. Kvantitativní odezva, jeden faktor

Jeden faktor o dvou úrovních: Srovnání dvou souborů dat, normální rozdělení

1. úroveň faktoru, n replikací: $X_1, X_2, \dots, X_n$ $X \sim N(\mu_X, \sigma_X^2)$
2. úroveň faktoru, m replikací: $Y_1, Y_2, \dots, Y_m$ $Y \sim N(\mu_Y, \sigma_Y^2)$

	1	2	3	...	n
A	x_1	x_2	x_3	...	x_n
B	y_1	y_2	y_3	...	y_m

- a) dvě nezávislá měření
- b) párová měření



2.2. Kvantitativní odezva, jeden faktor

Jeden faktor o dvou úrovních: Srovnání dvou souborů dat, normální rozdělení

a) dvě nezávislá měření:

1. úroveň faktoru, n replikací: $X_1, X_2, \dots, X_n$ $X \sim N(\mu_X, \sigma_X^2)$

2. úroveň faktoru, m replikací: $Y_1, Y_2, \dots, Y_m$ $Y \sim N(\mu_Y, \sigma_Y^2)$

	1	2	3	...	n
A	x_1	x_2	x_3	...	x_n
B	y_1	y_2	y_3	...	y_m

- oba parametry v obou případech známe
- známe střední hodnoty a neznáme rozptyly
- známe rozptyly a neznáme střední hodnoty
- žádný z parametrů neznáme

Odhady středních hodnot: $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ $\bar{Y} = \frac{1}{m} \sum_{i=1}^n Y_i$

Odhady rozptylů: $s_X^2 = \frac{1}{n-1} \sum (X - \bar{X})^2$, $s_Y^2 = \frac{1}{n-1} \sum (Y - \bar{Y})^2$



2.2. Kvantitativní odezva, jeden faktor

Jeden faktor o dvou úrovních: Srovnání dvou souborů dat, normální rozdělení

a) dvě nezávislá měření:

1. úroveň faktoru, n replikací: $X_1, X_2, \dots, X_n$ $X \sim N(\mu_X, \sigma_X^2)$

2. úroveň faktoru, m replikací: $Y_1, Y_2, \dots, Y_m$ $Y \sim N(\mu_Y, \sigma_Y^2)$

	1	2	3	...	n
A	x_1	x_2	x_3	...	x_n
B	y_1	y_2	y_3	...	y_m

- test shody rozptylů
- test shody středních hodnot při stejných rozptylech
- test shody středních hodnot při nestejných rozptylech



2.2. Srovnávací analýza: jeden faktor, dvě úrovně

Srovnání rozptylů dvou nezávislých měření - test shody rozptylů

Liší se statisticky významně dvě nezávislá měření z hlediska velikosti rozptylu?

Lze považovat rozptyl dvou nezávislých měření za shodný při dané hladině významnosti?

nulová hypotéza : $H_0 : \sigma_X^2 = \sigma_Y^2$

alternativní hypotéza: $H_A : \sigma_X^2 \neq \sigma_Y^2$

testová statistika : $F = \frac{s_X^2}{s_Y^2}$

hladina významnosti: α

F-test

Fisherovo-Snedecorovo
rozdělení $F(n-1, m-1)$

H_0 nezamítneme, když pro dané α bude

$$F_{\alpha/2}(n-1, m-1) < F < F_{\alpha/2}(n-1, m-1)$$

2.2. Srovnávací analýza

Srovnání středních hodnot dvou nezávislých souborů dat - dvouvýběrový t-test

Liší se statisticky významně dvě nezávislá měření z hlediska jejich střední hodnoty?

Lze považovat střední hodnoty dvou nezávislých měření za shodné při dané hladině významnosti?

Lze od sebe statisticky významně odlišit dvě nezávislá měření podle jejich střední hodnoty?

nulová hypotéza : $H_0 : \mu_X = \mu_Y$

alternativní hypotéza: $H_A : \mu_X \neq \mu_Y$ (oboustranná)

testová statistika : $T = \frac{\bar{X} - \bar{Y}}{s_{\bar{X} - \bar{Y}}}$

hladina významnosti: α

pokud $\sigma_X^2 = \sigma_Y^2$

dvouvýběrový t-test
se stejnými rozptyly

pokud $\sigma_X^2 \neq \sigma_Y^2$

dvouvýběrový t-test
s nesterýnými rozptyly

2.2. Srovnávací analýza

Srovnání středních hodnot dvou nezávislých souborů dat - dvouvýběrový t-test

Dvouvýběrový t-test se stejnými rozptyly:
$$T = \frac{\bar{X} - \bar{Y}}{s_{\bar{X} - \bar{Y}}}$$

pokud $\sigma_X^2 = \sigma_Y^2 = \sigma^2$

$$\Rightarrow s_{\bar{X} - \bar{Y}}^2 = s_{\bar{X}}^2 + s_{\bar{Y}}^2 = \frac{s_X^2}{n} + \frac{s_Y^2}{m} = \frac{s^2}{n} + \frac{s^2}{m} = s^2 \left(\frac{m + n}{m \cdot n} \right)$$

dále odhadneme s^2 ze všech naměřených hodnot:

$$s^2 = \frac{1}{n + m - 2} \left(\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^m (Y_i - \bar{Y})^2 \right) =$$
$$\frac{1}{n + m - 2} \left((n - 1)s_X^2 + (m - 1)s_Y^2 \right)$$

tedy:
$$s_{\bar{X} - \bar{Y}}^2 = \frac{n + m}{nm(n + m - 2)} \left((n - 1)s_X^2 + (m - 1)s_Y^2 \right)$$



2.2. Srovnávací analýza

Srovnání středních hodnot dvou nezávislých souborů dat - dvouvýběrový t-test

Dvouvýběrový t-test se stejnými rozptyly: $T = \frac{\bar{X} - \bar{Y}}{s_{\bar{X} - \bar{Y}}}$

$$s_{\bar{X} - \bar{Y}}^2 = \frac{n + m}{nm(n + m - 2)} \left((n - 1)s_X^2 + (m - 1)s_Y^2 \right)$$

Testová statistika bude mít tvar:

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{(n - 1)s_X^2 + (m - 1)s_Y^2}} \sqrt{\frac{nm(n + m - 2)}{n + m}}$$

ta má t-rozdělení (Studentovo rozdělení) pravděpodobnosti o $(n+m-2)$ stupních volnosti.

H_0 nezamítneme, když pro dané α bude $|T| \leq t_\alpha(n + m - 2)$ kde $t_\alpha(n + m - 2)$ je (oboustranná) α -kritická hodnota t-rozdělení o $(n+m-2)$ stupních volnosti.



2.2. Srovnávací analýza

Srovnání středních hodnot dvou nezávislých souborů dat - dvouvýběrový t-test

Dvouvýběrový t-test se nestejnými rozptyly: $T = \frac{\bar{X} - \bar{Y}}{s_{\bar{X} - \bar{Y}}}$

pokud $\sigma_X^2 \neq \sigma_Y^2$: $s_{\bar{X} - \bar{Y}}^2 = s_{\bar{X}}^2 + s_{\bar{Y}}^2 = \frac{s_X^2}{n} + \frac{s_Y^2}{m}$

Testová statistika bude mít tvar:

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{1}{n}s_X^2 + \frac{1}{m}s_Y^2}}$$

a má rozdělení, které je směsí t-rozdělení o $(n-1)$ a $(m-1)$ stupních volnosti.

H_0 nezamítneme, když pro dané α bude splněna nerovnost $|T| \leq At_\alpha(n-1) + Bt_\alpha(m-1)$, kde A a B jsou váhy, $A+B=1$.

$$A = \frac{\frac{1}{n}s_X^2}{\frac{1}{n}s_X^2 + \frac{1}{m}s_Y^2}, \quad B = \frac{\frac{1}{m}s_Y^2}{\frac{1}{n}s_X^2 + \frac{1}{m}s_Y^2}$$



2.2. Srovnávací analýza

Jeden faktor o dvou úrovních: Srovnání dvou souborů dat

b) párová pozorování:

1. úroveň faktoru, n replikací: $X_1, X_2, \dots, X_n$ $X \sim N(\mu_X, \sigma_X^2)$
2. úroveň faktoru, n replikací: $Y_1, Y_2, \dots, Y_n$ $Y \sim N(\mu_Y, \sigma_Y^2)$

	1	2	3	...	n
A	x_1	x_2	x_3	...	x_n
B	y_1	y_2	y_3	...	y_n

- pozorování stejné veličiny před a po nějakém zásahu
 - měření stejných objektů za různých podmínek
 - měření stejné veličiny ve dvou různých časech
 -
-
- párový test shody středních hodnot



2.2. Srovnávací analýza

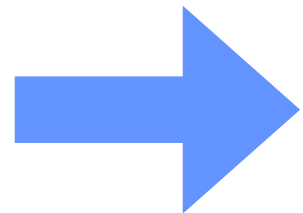
Srovnání párových souborů dat - párový t-test

1. úroveň faktoru, n replikací: $X_1, X_2, \dots, X_n$ $X \sim N(\mu_X, \sigma_X^2)$
2. úroveň faktoru, n replikací: $Y_1, Y_2, \dots, Y_n$ $Y \sim N(\mu_Y, \sigma_Y^2)$

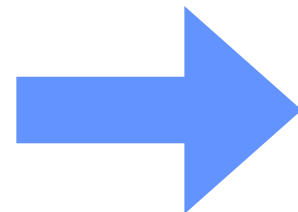
Testujeme $H_0 : \mu_X = \mu_Y$ proti $H_A : \mu_X \neq \mu_Y$

máme n párových pozorování : $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$

$$\Rightarrow Z_1 = X_1 - Y_1, Z_2 = X_2 - Y_2, \dots, Z_n = X_n - Y_n,$$
$$Z \sim N(\mu_X - \mu_Y, \sigma_Z^2)$$



Testujeme $H_0 : \mu_Z = 0$ proti $H_A : \mu_Z \neq 0$



Jednovýběrový t-test o střední hodnotě veličiny Z



2.2. Srovnávací analýza

Srovnání párových souborů dat - párový t-test

Jednovýběrový t-test :

$$T = \frac{\bar{Z} - a}{s_Z} \sqrt{n}$$

$$H_0 : \mu_Z = a$$

$$H_A : \mu_Z \neq a$$

T má t-rozdělení (Studentovo rozdělení) pravděpodobnosti o $(n-1)$ stupních volnosti.

H_0 nezamítneme, když pro dané α bude $|T| \leq t_\alpha(n-1)$ kde $t_\alpha(n-1)$ je (oboustranná) α -kritická hodnota t-rozdělení o $(n-1)$ stupních volnosti.

2.2. Srovnávací analýza

Jednostranné srovnávací testy

“dolní” nebo “horní” jednostranná alternativa :

$$H_0 : \mu_X = \mu_Y$$

$$H_A : \mu_X < \mu_Y$$

$$H_0 : \mu_X = \mu_Y$$

$$H_A : \mu_X > \mu_Y$$

H_0 nezamítneme, když pro dané α bude buď $T > t_\alpha(n - 1)$ nebo $T < -t_\alpha(n - 1)$, kde $t_\alpha(n - 1)$ je (jednostranná) α -kritická hodnota t-rozdělení o $(n-1)$ stupních volnosti.

oboustranná α -kritická hodnota je $(1 - \alpha/2)$ -kvantil $t_{1-\alpha/2}(n - 1)$

jednostranná α -kritická hodnota je $(1 - \alpha)$ -kvantil $t_{1-\alpha}(n - 1)$

2.2. Srovnávací analýza

Příklady:

Lze považovat délky tyčí od dvou různých dodavatelů za shodné na hladině významnosti 5%?

Byly měřeny odchylky délky ocelových tyčí od požadované hodnoty 4m od dvou dodavatelů. Odchylky jsou uvedeny v cm.

Dodavatel X:

```
> x
[1] 0.41379418 0.51040227 3.28722973 7.31995568 4.53994434 -1.07426821
[7] 4.74575978 2.55201407 3.22058685 -1.17401554 -1.24119500 4.18294690
[13] 0.65486399 -0.18908709 -0.73101186 1.27876451 1.26734875 2.78570344
[19] 2.96834139 1.22145702 1.80851440 -0.80356569 2.57347292 3.42552806
[25] 1.66904559 -2.21179295 4.17696270 2.15191523 3.62707736 0.06900211
[31] 0.51371315 0.54983237 4.09554316 1.28465289 4.05350899 5.10504379
[37] 4.25580572 0.79826235 -1.02042629 1.87299786 0.14051938 3.05622839
[43] 4.74780021 4.54794140 -6.54132331 1.94429658 1.95488616 4.73267571
[49] 4.83082378 2.95830720 2.99769818 -1.07337799 0.58403864 2.73050678
[55] 0.28021230 10.49771713 2.36870296 0.60689702 8.42679434 1.29763889
[61] 1.31289734 1.93230073 5.92597773 1.49746935 6.30721756 3.15585521
[67] 5.38824907 3.27322441 3.41248356 -0.40437473 3.19350142 -4.06261001
[73] -1.05763312 -0.39748962 0.86637433 2.02108109 -1.06445976 1.10375263
[79] 4.51823259 -0.75725877 -0.87173075 -2.19932463 7.70167909 1.48655986
[85] 4.90757730 5.51652338 -0.34615559 0.01031344 4.57582354 1.17516968
[91] -0.21932558 -1.27848277 2.97655676 1.44863955 3.67881403 0.30868429
[97] -2.52052309 0.05248743 0.07728483 -1.12975005 3.99585182 0.79045260
[103] 3.73159608 7.36490361 6.40646375 -1.54228149 -0.65100869 4.04305846
[109] 2.47766853 -3.48957597 6.20840771 0.40560482 0.49118447 -1.48277951
[115] -1.23675030 5.16138353 1.15383008 2.75286404 4.70183189 -2.29877355
```

2.2. Srovnávací analýza

Příklady:

Lze považovat délky tyčí od dvou různých dodavatelů za shodné na hladině významnosti 5%?

Byly měřeny odchylky délky ocelových tyčí od požadované hodnoty 4m od dvou dodavatelů. Odchylky jsou uvedeny v cm.

Dodavatel Y:

```
> y
[1] 6.65956934 2.78876119 0.33397602 -0.03763918 0.74993937 3.81490677
[7] 1.70428804 -3.31291341 -0.22972370 4.02124752 5.93229834 3.30506070
[13] -3.61277063 0.78809415 0.37976841 1.52357320 1.76230055 1.03078642
[19] -2.74093726 2.77205578 -0.25596771 -0.79295335 -1.99567925 7.14183490
[25] 6.56129569 -2.39785588 -2.30807391 -1.02088455 -2.26040839 -2.76088135
[31] 1.81877126 0.14669279 4.21783231 -2.13184320 3.69196005 -2.69614367
[37] -2.68014820 3.72209577 1.73709472 -0.70580812 0.07337669 2.17063230
[43] 2.72495294 5.04390706 1.32219033 4.72349163 -0.67638087 2.64424944
[49] 2.78769261 -2.10997705 4.26042721 -3.50266144 1.72564280 -2.07028305
[55] -4.59779260 -1.71953774 2.90307934 1.38358058 3.42339203 -1.68000430
[61] 7.55683608 6.32574310 -2.60318964 3.24511198 0.97390332 2.22611398
[67] 0.83831831 0.07828888 2.29402602 2.68356827 0.07483911 3.38214384
[73] -0.59180508 9.07209729 -1.27708114 4.77997853 -0.83918672 6.26383807
[79] 1.50674691 3.25716693 5.70351834 5.80174051 3.61099316 2.19293272
[85] -1.46102337 -0.97135778 1.54849399 4.34257358 -1.64886246 2.44942102
[91] 2.68469434 1.64707956 5.49827517 1.01640668 4.43099277 2.23430799
[97] -1.74337571 6.43458332 2.94137432 -1.01569579
```

2.2. Srovnávací analýza

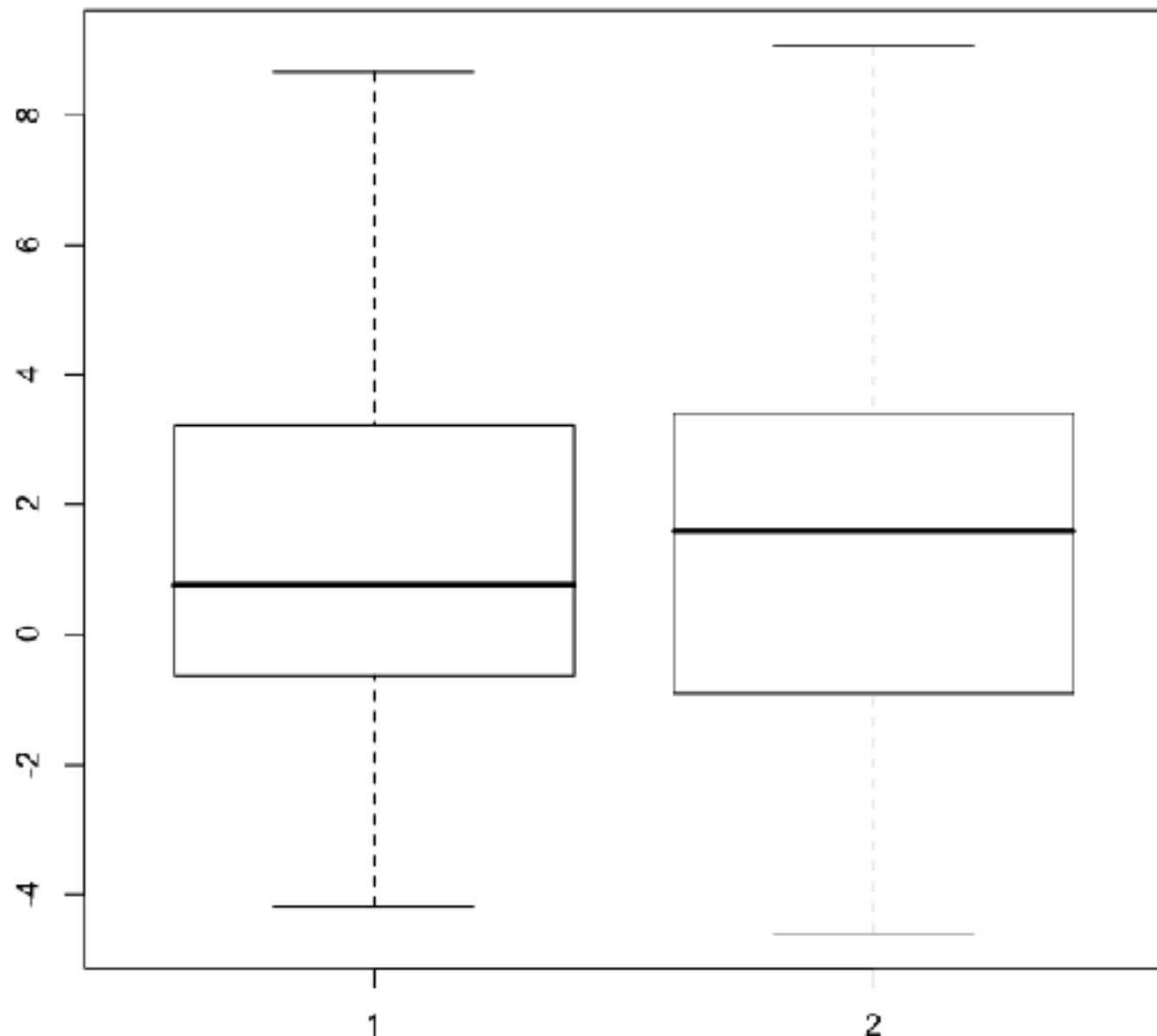
Příklady:

Lze považovat délky tyčí od dvou různých dodavatelů za shodné na hladině významnosti 5%?

Byly měřeny odchylky délky ocelových tyčí od požadované hodnoty 4m od dvou dodavatelů. Odchylky jsou uvedeny v cm.

1) Vizualizace dat:

Box&Whiskers diagram



2.2. Srovnávací analýza

Příklady:

Lze považovat délky tyčí od dvou různých dodavatelů za shodné na hladině významnosti 5%?

Byly měřeny odchylky délky ocelových tyčí od požadované hodnoty 4m od dvou dodavatelů. Odchylky jsou uvedeny v cm.

2) Srovnání rozptylů: F-test

```
> var.test(x,y)

F test to compare two variances

data:  x and y
F = 0.8712, num df = 119, denom df = 99, p-value = 0.4701
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.5943383 1.2684711
sample estimates:
ratio of variances
      0.8711758
```

=> nulovou hypotézu nezamítáme, rozptyly se statisticky významně neliší

2.2. Srovnávací analýza

Příklady:

Lze považovat délky tyčí od dvou různých dodavatelů za shodné na hladině významnosti 5%?

Byly měřeny odchylky délky ocelových tyčí od požadované hodnoty 4m od dvou dodavatelů. Odchylky jsou uvedeny v cm.

3) Srovnání středních hodnot: dvouvýběrový t-test se shodnými rozptyly

```
> t.test(x,y, var.equal=T)
```

```
Two Sample t-test
```

```
data: x and y
```

```
t = 1.0375, df = 218, p-value = 0.3007
```

```
alternative hypothesis: true difference in means is not equal to 0
```

```
95 percent confidence interval:
```

```
-0.3598731  1.1598308
```

```
sample estimates:
```

```
mean of x mean of y
```

```
1.884360  1.484381
```

=> nulovou hypotézu nezamítáme, střední hodnoty se statisticky významně neliší

2.2. Srovnávací analýza

Příklady:

Bylo zjišťováno, zda množství vyrobených kusů za hodinu (výkonnost) závisí na typu šicího stroje.

- Odezva: výkon vyjádřený počtem výrobků za hodinu
- Faktory: typ stroje (2 úrovně: se servomotorem, s asynchronním motorem)
- Bloky: 8 operátorů, každý bude pracovat na obou typech stroje

Jedná se o jedno-faktorový experiment se 2 úrovněmi faktoru se závislými měřeními
=> blokové (párové) uspořádání experimentu, k vyhodnocení použijeme párový t-test

1. Data:

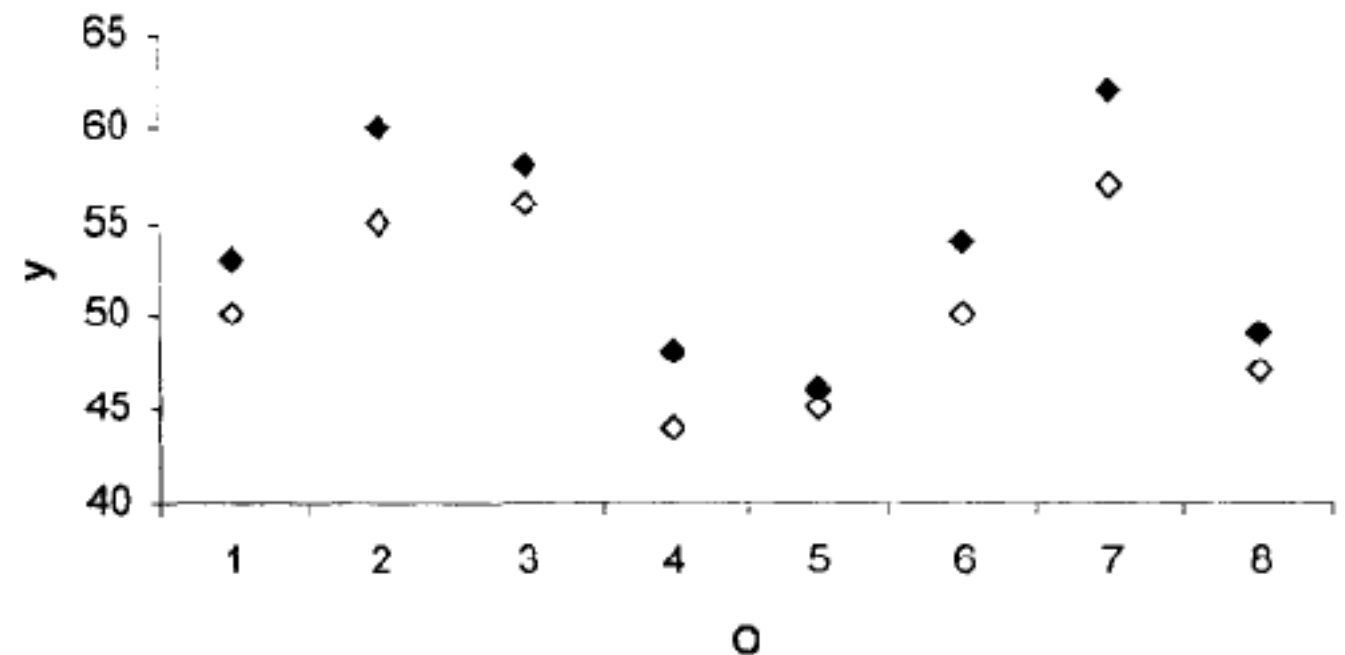
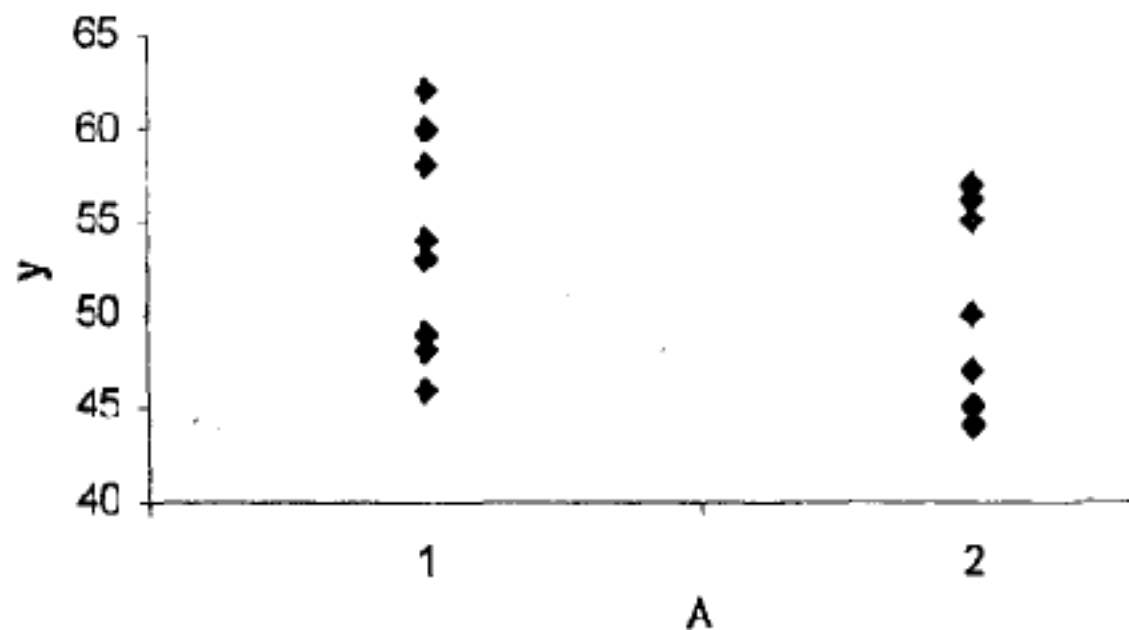
operátor	1	2	3	4	5	6	7	8
stroj A1	53	60	58	48	46	54	62	49
stroj A2	50	55	56	44	45	50	57	47
rozdíl dj	3	5	2	4	1	4	5	2

2.2. Srovnávací analýza

Příklady:

Bylo zjišťováno, zda množství vyrobených kusů za hodinu (výkonnost) závisí na typu šicího stroje.

2. Grafické znázornění dat:



2.2. Srovnávací analýza

Příklady:

Bylo zjišťováno, zda množství vyrobených kusů za hodinu (výkonnost) závisí na typu šicího stroje.

3. Párový t-test:

$$H_0 : \mu_{A1} = \mu_{A2}, \quad H_A : \mu_{A1} \neq \mu_{A2}$$

$$T = \frac{\bar{d}}{\sqrt{\frac{s_d^2}{r}}}$$

kde

$$\bar{d} = \frac{1}{r} \sum_{j=1}^r d_j, \quad s_d^2 = \frac{1}{r-1} \sum_{j=1}^r (d_j - \bar{d})^2$$

	y1	y2
Stř. hodnota	53,75	50,5
Rozptyl	34,5	25,42857
Pozorování	8	8
Pears. korelace	0,974278	
Hyp. rozdíl stř. hodnot	0	
Stupně volnosti	7	
t stat	6,177483	
P(T<=t) (1)	0,000228	
t krit (1)	1,894578	
P(T<=t) (2)	0,000455	
t krit (2)	2,364623	

=> nulovou hypotézu zamítáme, typ stroje ovlivňuje výkonnost, která je statisticky významně vyšší u prvního typu stroje

2.2. Srovnávací analýza

Příklady:

Byla měřena rychlost reakce operátorů před a po speciálním cvičení v sekundách. Mělo cvičení statisticky významný vliv na rychlost?

1) Data:

```
> pred_cviceni
[1] 12.666378  7.322789 15.021706 13.616913 10.970712  5.464451
[7]  9.999636 15.693764 13.771444 17.065310  6.940708 15.860749
[13] 18.019348  6.326531 20.647763 23.005369 14.619170 20.787108
[19] 14.238225  9.674337 14.763170  9.613791  9.727326  9.146292
[25] 21.246960 16.200128 15.466065 13.691879  9.032113 10.558392
[31] 18.258896 14.992416 14.722569 10.579842 10.758363  8.894299
[37] 13.502299 12.994734 14.775563  9.818535 18.208089  8.438143
[43]  8.282819 11.090392 15.174881  7.704479  8.917742 10.275903
[49] 11.488700 16.572150 18.892428 13.544225  9.309845 13.713258
[55] 12.904993  8.951567  9.041688 10.222305 14.136072  9.222289
[61] 15.208694 14.627659 15.287092 11.389052  7.716052 14.307632
[67] 14.647653 18.705963 13.665201  8.025347 13.157791 14.336731
[73]  9.548584 12.522605 11.876452 12.241549 12.944160 17.637175
[79]  9.854223 17.877400 15.892081  9.893356  7.791175 11.901961
[85] 15.605362 13.464186 12.451922 16.090626  8.907932 16.333859
[91] 13.554146 19.586575 11.765020  9.981692  5.325750 20.168371
[97] 12.485393 14.349888 14.198229  7.315012 16.787920 10.998550
[103] 10.377856 13.531181 12.258939 11.346062 12.998020  8.498104
[109] 14.195263 15.372914 11.698431 12.929311 11.232474 21.551867
[115] 10.436798 14.430260 18.836296 14.838428 14.450987 10.879682
```

2.2. Srovnávací analýza

Příklady:

Byla měřena rychlost reakce operátorů před a po speciálním cvičení v sekundách. Mělo cvičení statisticky významný vliv na rychlost?

1) Data:

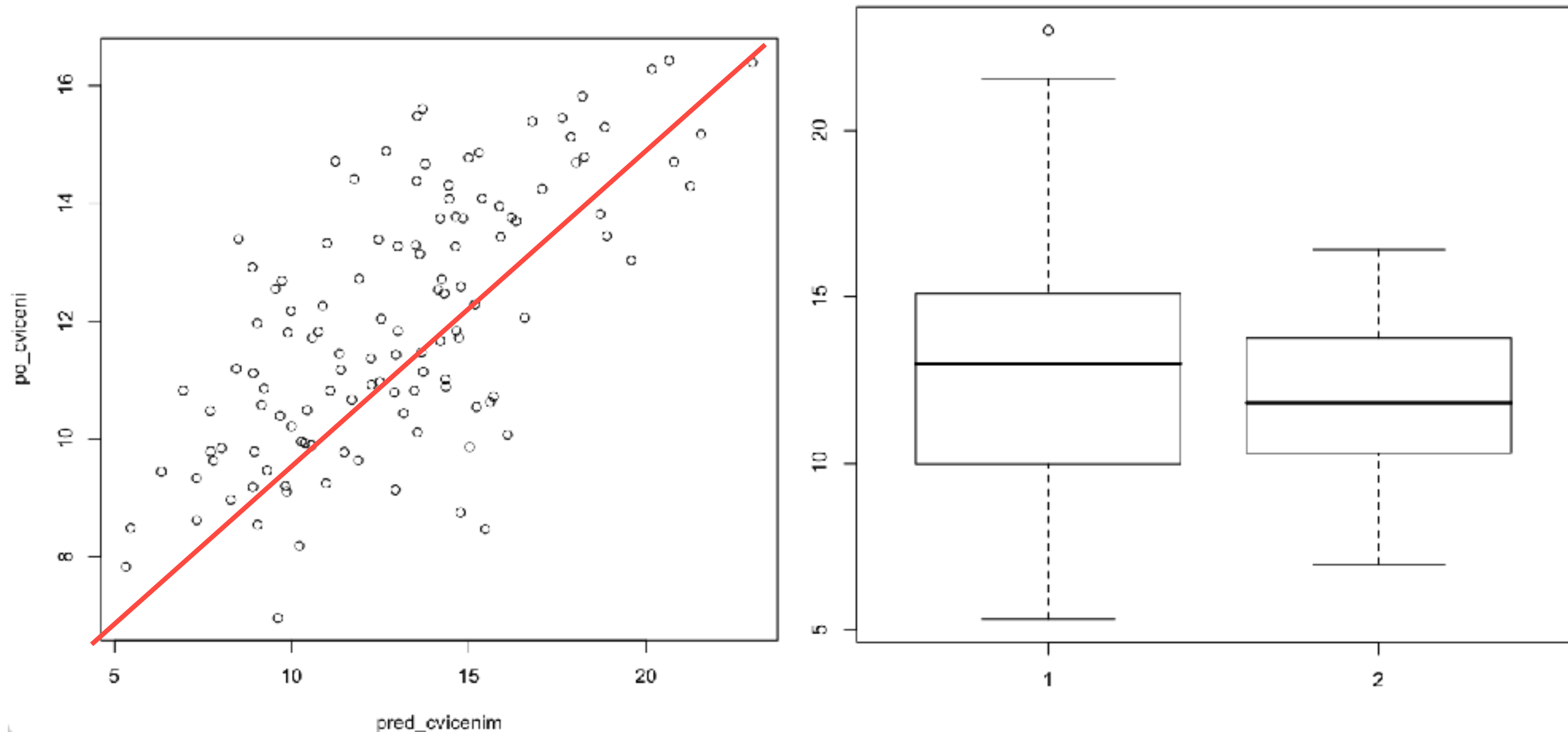
```
> po_cviceni
[1] 14.889379  8.627612  9.867455 13.141168  9.249122  8.490774
[7] 10.217290 10.724403 14.669450 14.243944 10.826905 13.951521
[13] 14.693401  9.449562 16.425888 16.392689 13.265474 14.704994
[19] 12.718107 10.395385  8.756276  6.961521 12.688497 10.578342
[25] 14.294064 13.763032  8.472324 15.605253 11.968936  9.897284
[31] 14.788205 14.773378 11.723336 11.719464 11.824407 12.914485
[37] 13.291805 13.272867 12.586791  9.202608 15.817188 11.197137
[43]  8.974410 10.823942 12.289400 10.483861 11.119684  9.956822
[49]  9.778551 12.062084 13.449972 15.481139  9.470557 11.143402
[55] 10.793291  9.786869  8.547580  8.188947 12.532635 10.862473
[61] 10.547040 13.774638 14.861969 11.180668  9.790466 12.469556
[67] 11.837173 13.820717 11.476120  9.850563 10.440890 11.015557
[73] 12.547672 12.041457  9.639740 11.368657 11.431948 15.449064
[79]  9.110052 15.125478 13.433802 11.807514  9.632299 12.725762
[85] 10.628523 10.824474 13.389953 10.077884  9.185360 13.697777
[91] 10.116078 13.036067 14.412094 12.175099  7.835201 16.277825
[97] 10.967441 10.892966 11.668289  9.340267 15.392018 13.323701
[103]  9.928631 14.378075 10.924935 11.448320 11.836161 13.397990
[109] 13.744963 14.083459 10.668370  9.139692 14.716621 15.173684
[115] 10.493444 14.308470 15.295041 13.748886 14.074436 12.261138
```

2.2. Srovnávací analýza

Příklady:

Byla měřena rychlost reakce operátorů před a po speciálním cvičení v sekundách. Mělo cvičení statisticky významný vliv na rychlost?

2) Grafické zobrazení



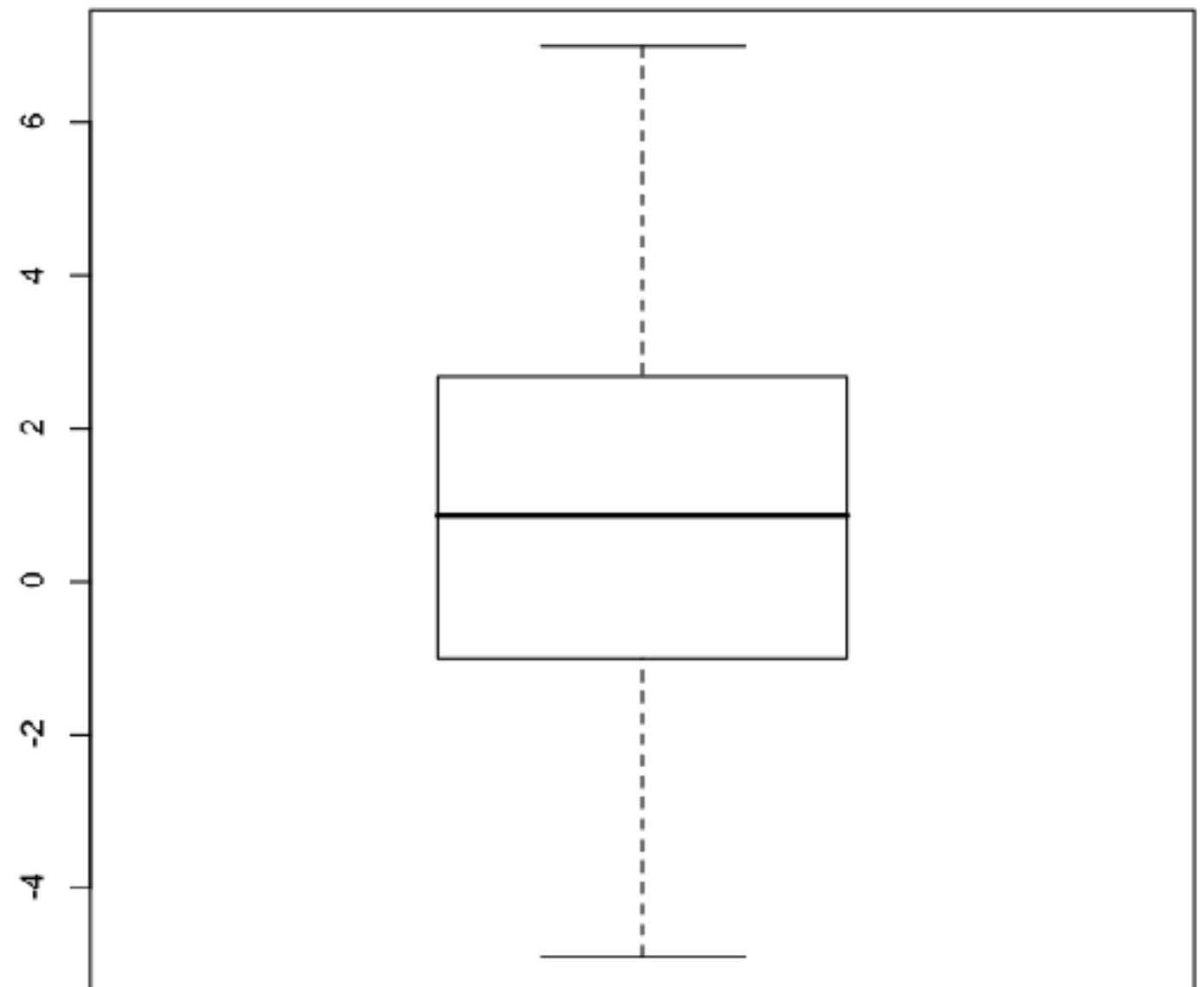
2.2. Srovnávací analýza

Příklady:

Byla měřena rychlost reakce operátorů před a po speciálním cvičení v sekundách. Mělo cvičení statisticky významný vliv na rychlost?

3) Rozdíly:

```
> rozdil = pred_cviceni - po_cviceni
> rozdil
 [1] -2.22300091 -1.30482283  5.1
 [7] -0.21765415  4.96936015 -0.8
[13]  3.32594695 -3.12303070  4.2
[19]  1.52011877 -0.72104757  6.0
[25]  6.95289576  2.43709534  6.9
[31]  3.47069129  0.21903888  2.9
[37]  0.21049440 -0.27813376  2.1
[43] -0.69159043  0.26644963  2.8
[49]  1.71014906  4.51006508  5.4
[55]  2.11170209 -0.83530141  0.4
[61]  4.66165463  0.85302053  0.4
[67]  2.81048012  4.88524631  2.1
[73] -2.99908758  0.48114855  2.2
[79]  0.74417071  2.75192209  2.4
[85]  4.97683829  2.63971186 -0.9
[91]  3.43806805  6.55050838 -2.6
[97]  1.51795203  3.45692228  2.5
[103]  0.44922574 -0.84689404  1.3
[109]  0.45029949  1.28945495  1.0
[115] -0.05664624  0.12179055  3.5
```



2.2. Srovnávací analýza

Příklady:

Byla měřena rychlost reakce operátorů před a po speciálním cvičení v sekundách. Mělo cvičení statisticky významný vliv na rychlost?

4) Párový t-test:

```
> t.test(rozdil, mu=0)

One Sample t-test

data:  rozdil
t = 4.0391, df = 119, p-value = 9.54e-05
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.5089508 1.4878397
sample estimates:
mean of x
0.9983952
```

=> nulovou hypotézu zamítáme, cvičení mělo vliv a rychlost reakce se statisticky významně zvýšila

2.2. Kvantitativní odezva, jeden faktor

Jeden faktor o dvou úrovních: Srovnání dvou souborů dat, obecné rozdělení

1. úroveň faktoru, n replikací: $X_1, X_2, \dots, X_n$

2. úroveň faktoru, m replikací: $Y_1, Y_2, \dots, Y_m$

a) dvě nezávislá měření

b) párová měření

	1	2	3	...	n
A	x_1	x_2	x_3	...	x_n
B	y_1	y_2	y_3	...	y_m

Neparametrické testy pro srovnání dvou souborů dat:

1) Dvouvýběrový Wilcoxonův test

2) Párový (jednovýběrový) Wilcoxonův test

3) Znaménkový test



2.2. Kvantitativní odezva, jeden faktor

Jeden faktor o dvou úrovních: Srovnání dvou souborů dat, obecné rozdělení

Neparametrické testy jsou založeny na **pořadových statistikách**:

Dvouvýběrový Wilcoxonův test $H_0 : \mu_X = \mu_Y$ $H_A : \mu_X \neq \mu_Y$

- Sloučíme všechna měření v jeden soubor: $X_1, X_2, \dots, X_n, Y_1, Y_2, \dots, Y_m$
- Uspořádáme je podle velikosti a potom určíme:
 - R_k^X = pořadí k-tého pozorování X ve spojeném souboru
 - R_l^Y = pořadí l-tého pozorování Y ve spojeném souboru
 - R^X = součet všech R_k^X pro $k=1, \dots, n$ a R^Y = součet všech R_l^Y pro $l=1, \dots, m$

Pro malá n, m (do 30) porovnáváme přímo jednu z hodnot buď R^X nebo R^Y s kritickými hodnotami pro dvouvýběrový Wilcoxonův test (viz např.

https://is.muni.cz/th/r5oe7/Bakalarska_prace.pdf), pro velké hodnoty n a m (nad 30) lze jako testovou statistiku použít

$$W = \frac{R^X - E(R^X)}{\sqrt{\text{var}(R^X)}}$$

$$\text{kde } E(R^X) = \frac{n(m+n+1)}{2}, \quad \text{var}(R^X) = \frac{mn(m+n+1)}{12}$$

Tato statistika má přibližně $N(0,1)$ rozdělení.



2.2. Kvantitativní odezva, jeden faktor

Jeden faktor o dvou úrovních: Srovnání dvou souborů dat, obecné rozdělení

Neparametrické testy jsou založeny na **pořadových statistikách**:

Párový Wilcoxonův test

- Spočítáme rozdíly $(X_1 - Y_1), (X_2 - Y_2), \dots, (X_n - Y_n)$ a vyloučíme nulové rozdíly
- Uspořádáme je podle velikosti absolutních hodnot a potom určíme:
 - R_k^+ = pořadí k-tého kladného rozdílu, R_l^- = pořadí l-tého záporného rozdílu
 - $R^+ =$ součet všech R_k^+ a $R^- =$ součet všech R_l^-

Pro malá n, m (do 30) porovnáváme přímo jednu z hodnot buď R^+ nebo R^- s kritickými hodnotami pro jednovýběrový Wilcoxonův test (viz např. https://is.muni.cz/th/r5oe7/Bakalarska_prace.pdf), pro velké hodnoty n a m (nad 30) lze jako testovou statistiku použít

$$V = \frac{R^+ - E(R^+)}{\sqrt{\text{var}(R^+)}}$$

kde $E(R^+) = \frac{n^*(n^* + 1)}{4}$, $\text{var}(R^+) = \frac{n^*(n^* + 1)(2n^* + 1)}{24}$ a n^* je počet nenulových rozdílů. Tato statistika má přibližně $N(0,1)$ rozdělení.



2.2. Kvantitativní odezva, jeden faktor

Jeden faktor o dvou úrovních: Srovnání dvou souborů dat, obecné rozdělení

Neparametrické testy jsou založeny na **pořadových statistikách**:

Znaménkový test (párový test)

- Spočítáme rozdíly $(X_1 - Y_1), (X_2 - Y_2), \dots, (X_n - Y_n)$ a vyloučíme nulové rozdíly
- Spočteme počet záporných rozdílů Z a počet nenulových rozdílů n^*

Statistika Z má binomické rozdělení s parametry n^* a $1/2$, pro velké hodnoty n^* ji lze aproximovat normálním rozdělením a tedy test založit na statistice

$$S = \frac{|2Z - n^*|}{\sqrt{n^*}}$$

která má přibližně $N(0,1)$ rozdělení. Hypotézu o shodě zamítáme, když je $S \geq u(1 - \alpha/2)$.

