

# Základy stochastiky

## X. Úvod do matematické statistiky

Prof. RNDr. Gejza Dohnal, CSc.

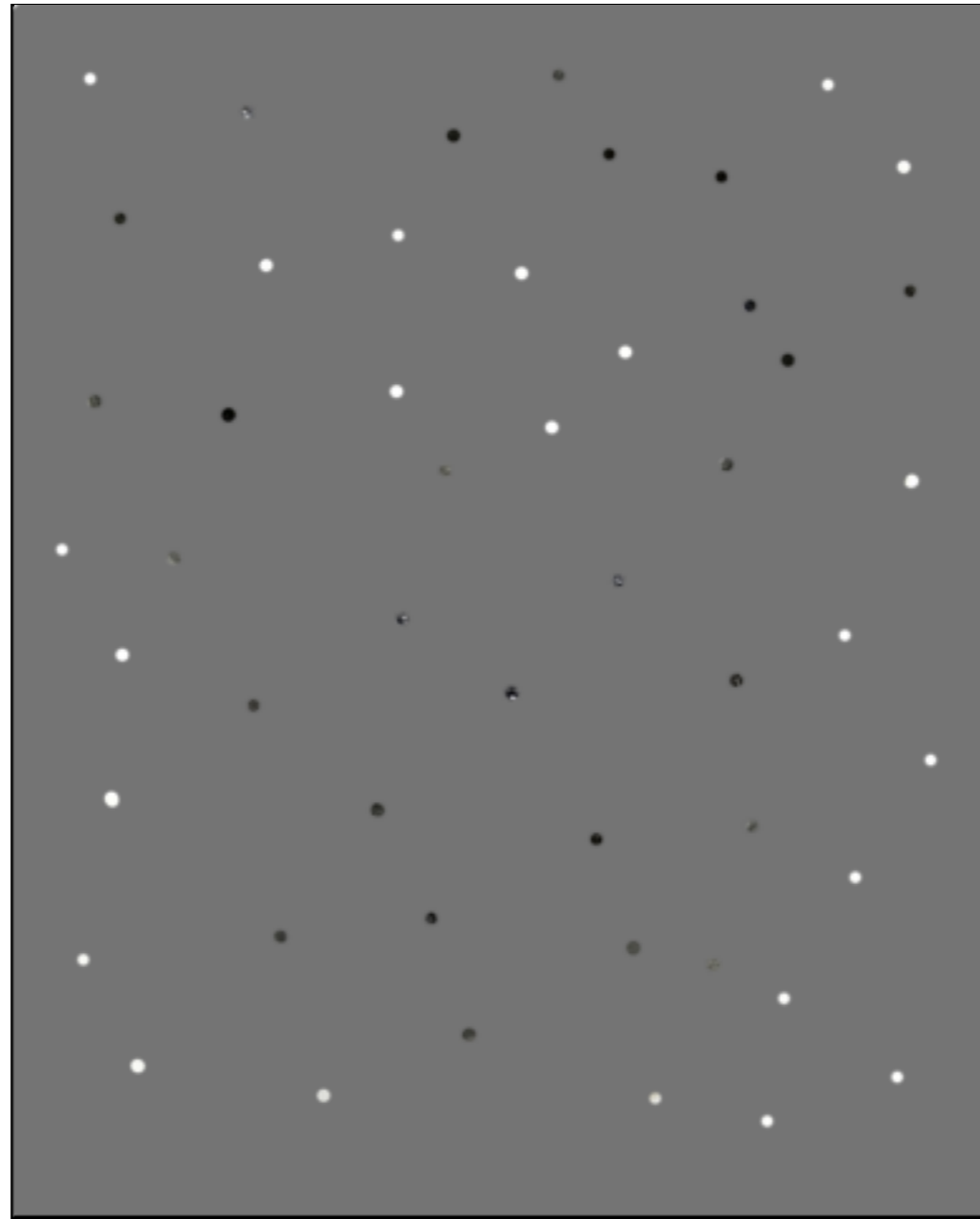


# Úloha statistické indukce



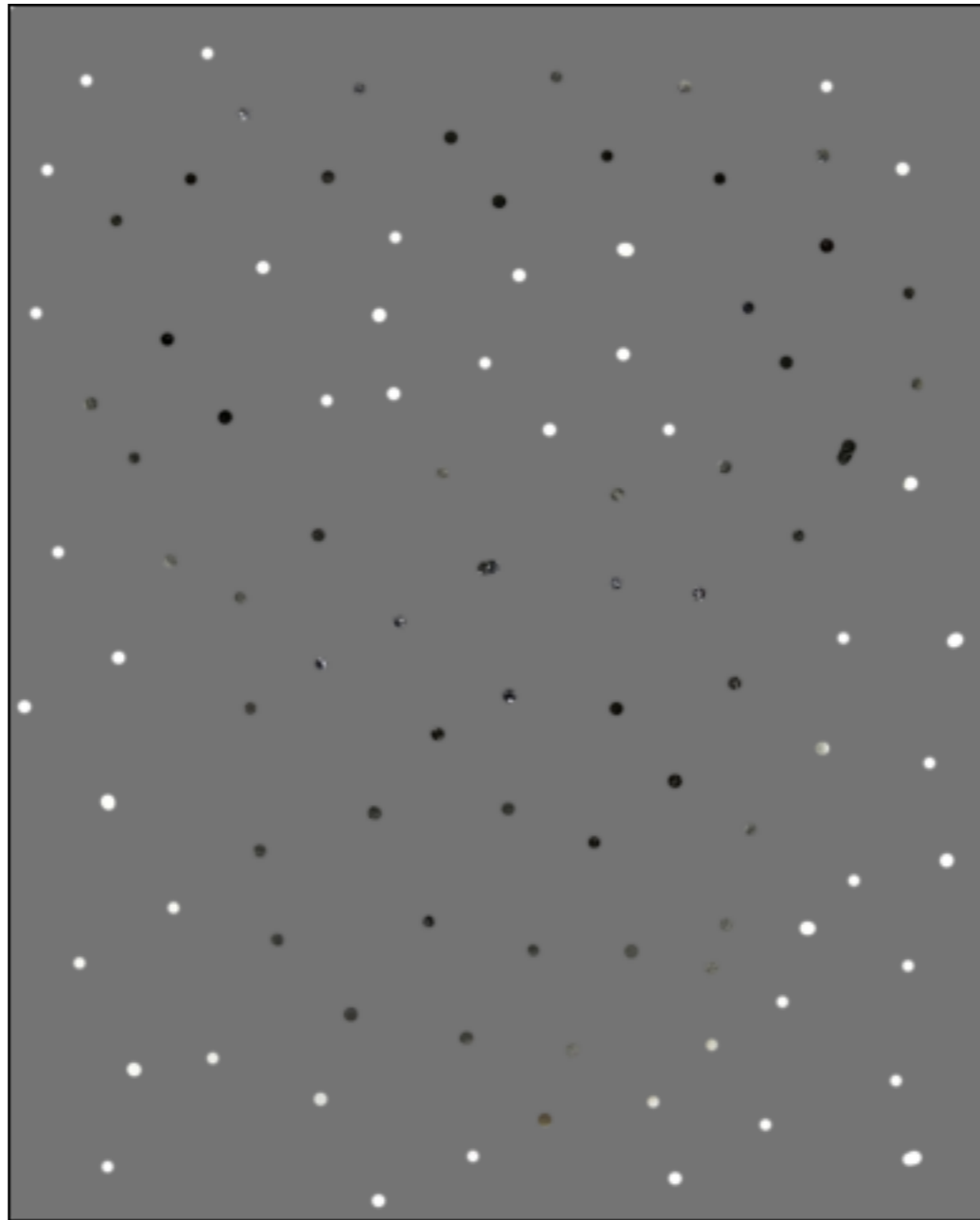
# Úloha statistické indukce

$n=50$



# Úloha statistické indukce

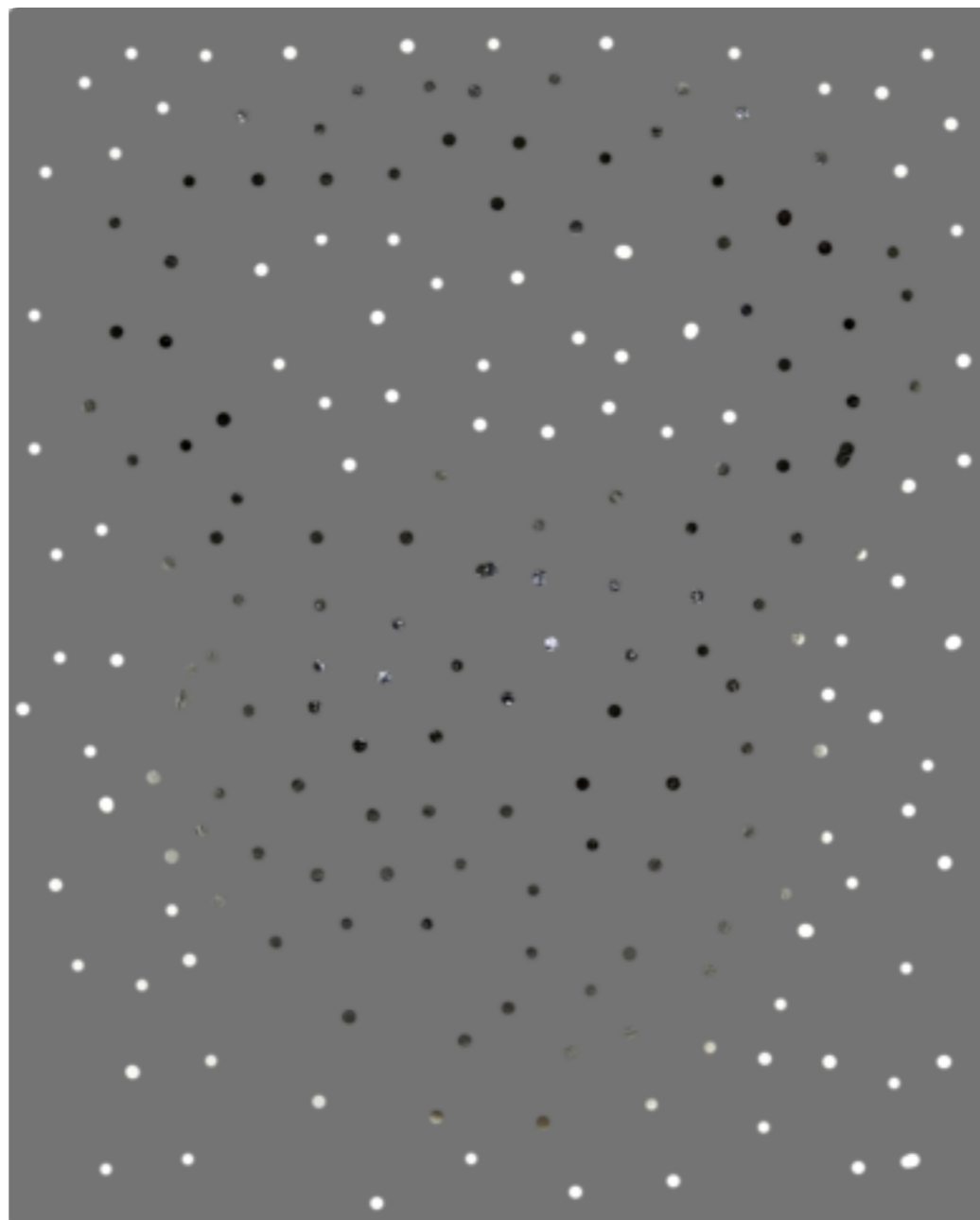
$n=100$





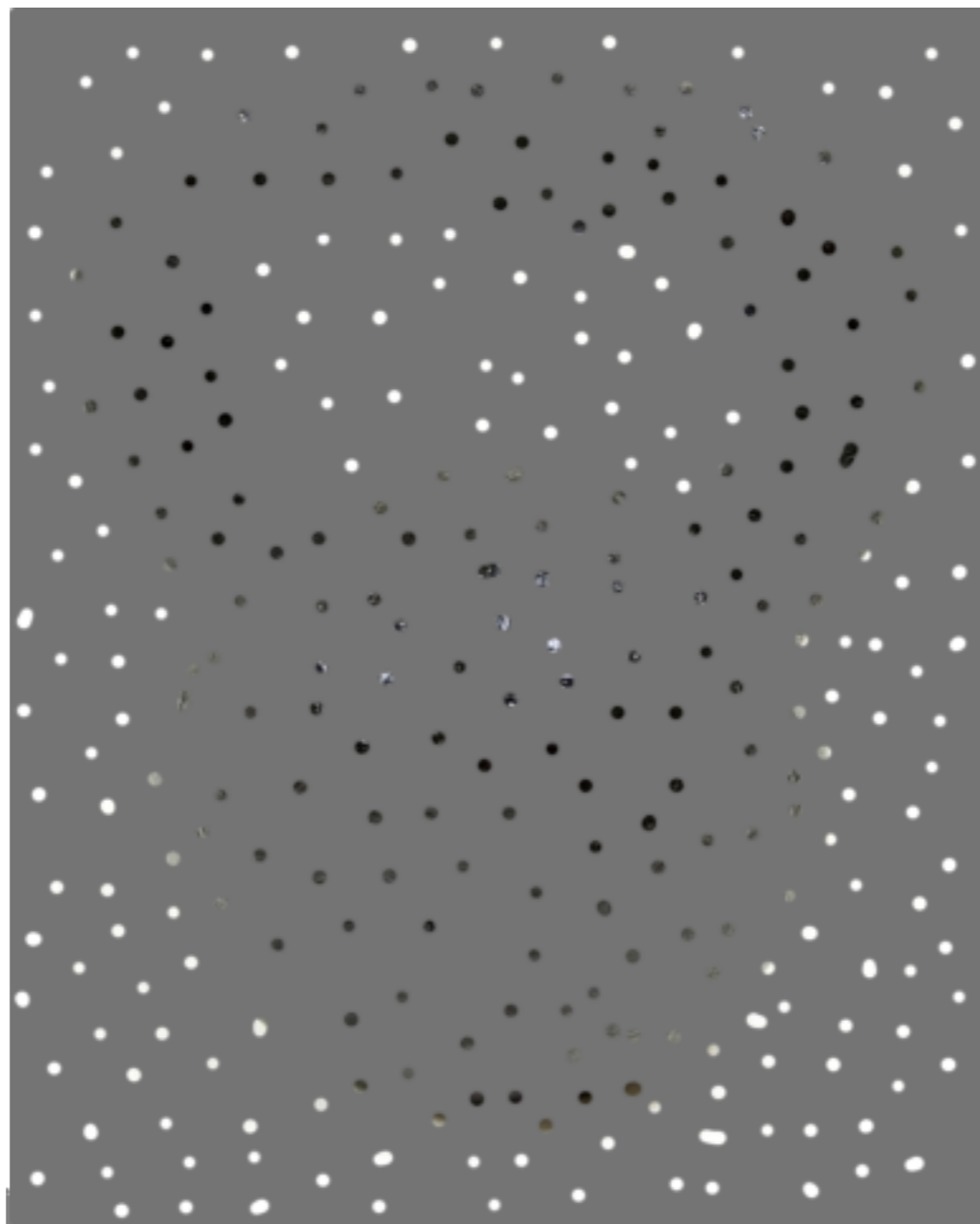
# Úloha statistické indukce

$n=200$

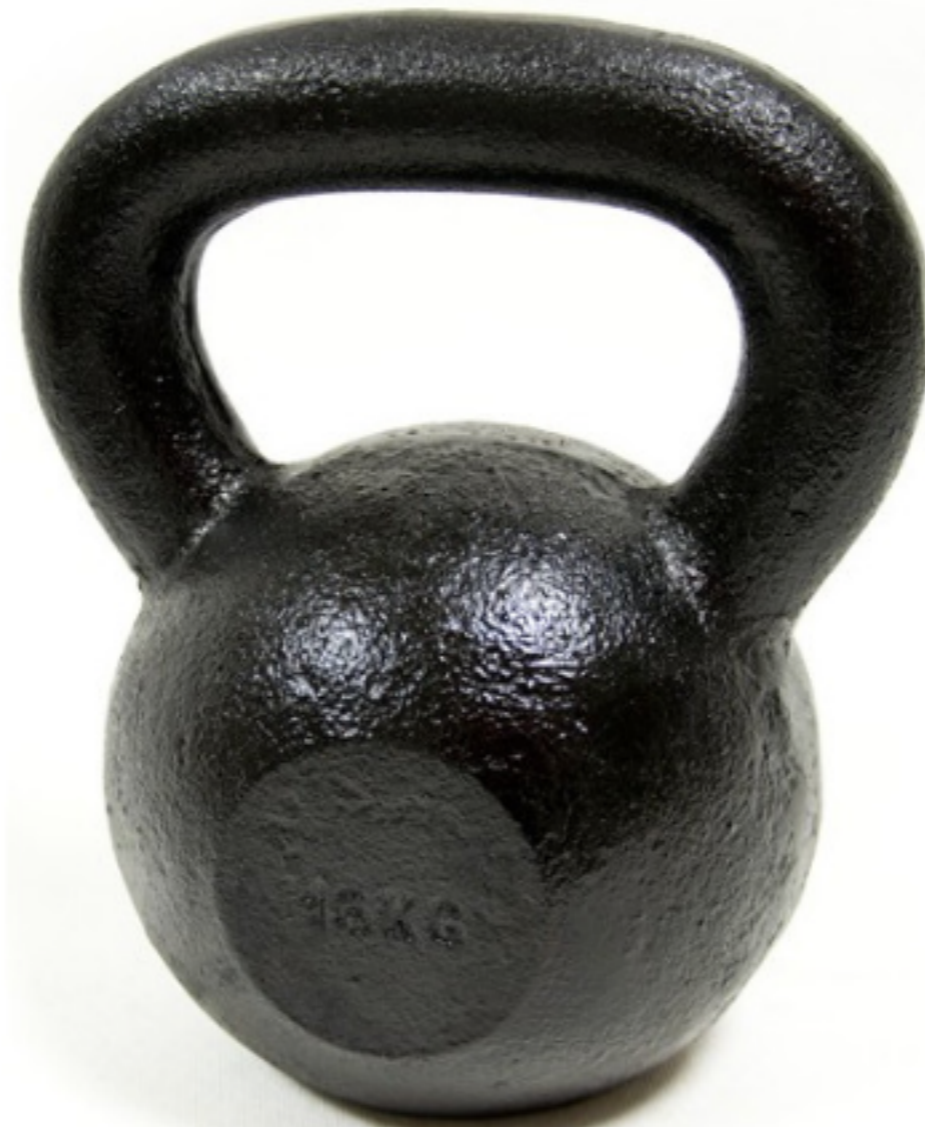


# Úloha statistické indukce

$n=300$

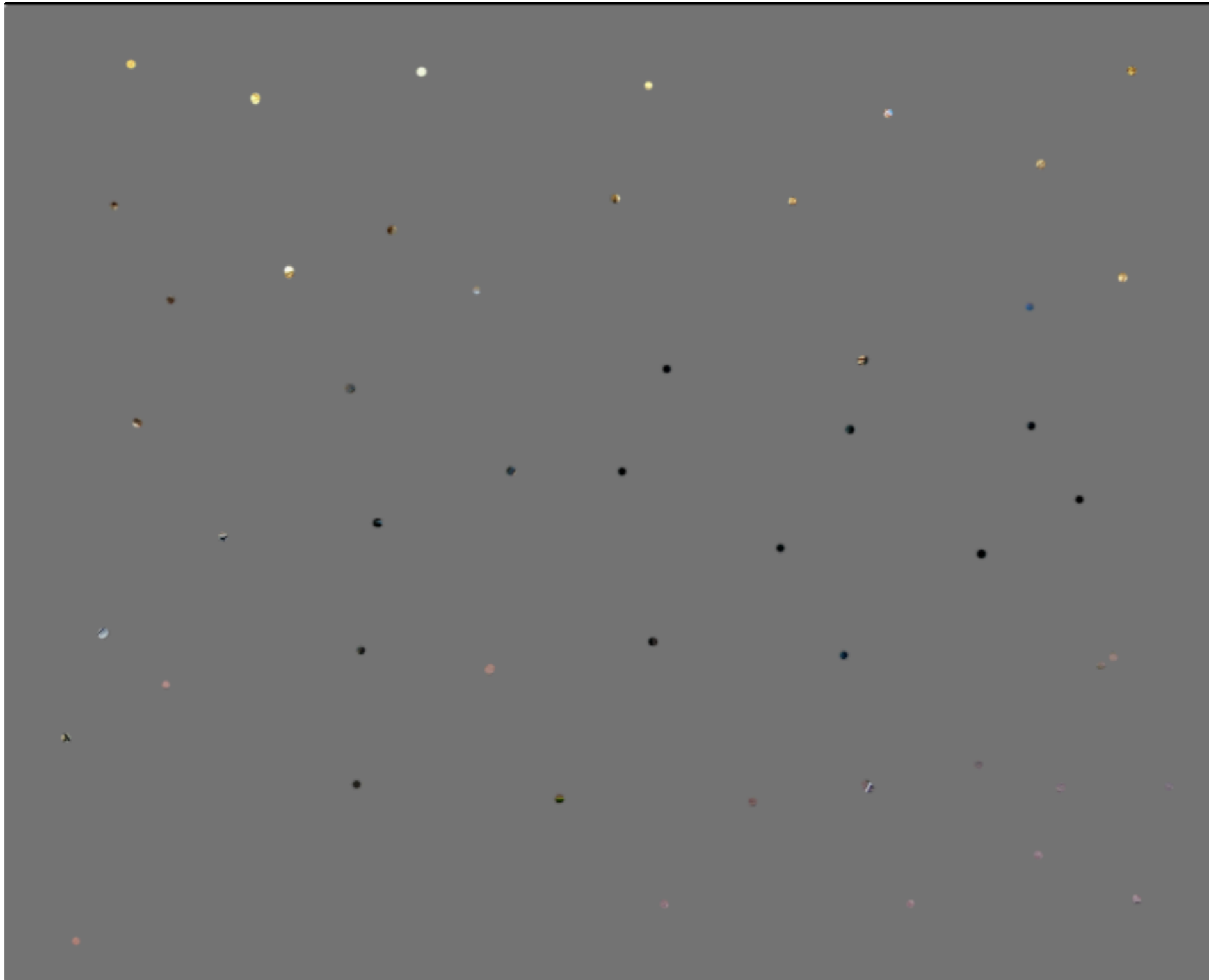


# Úloha statistické indukce



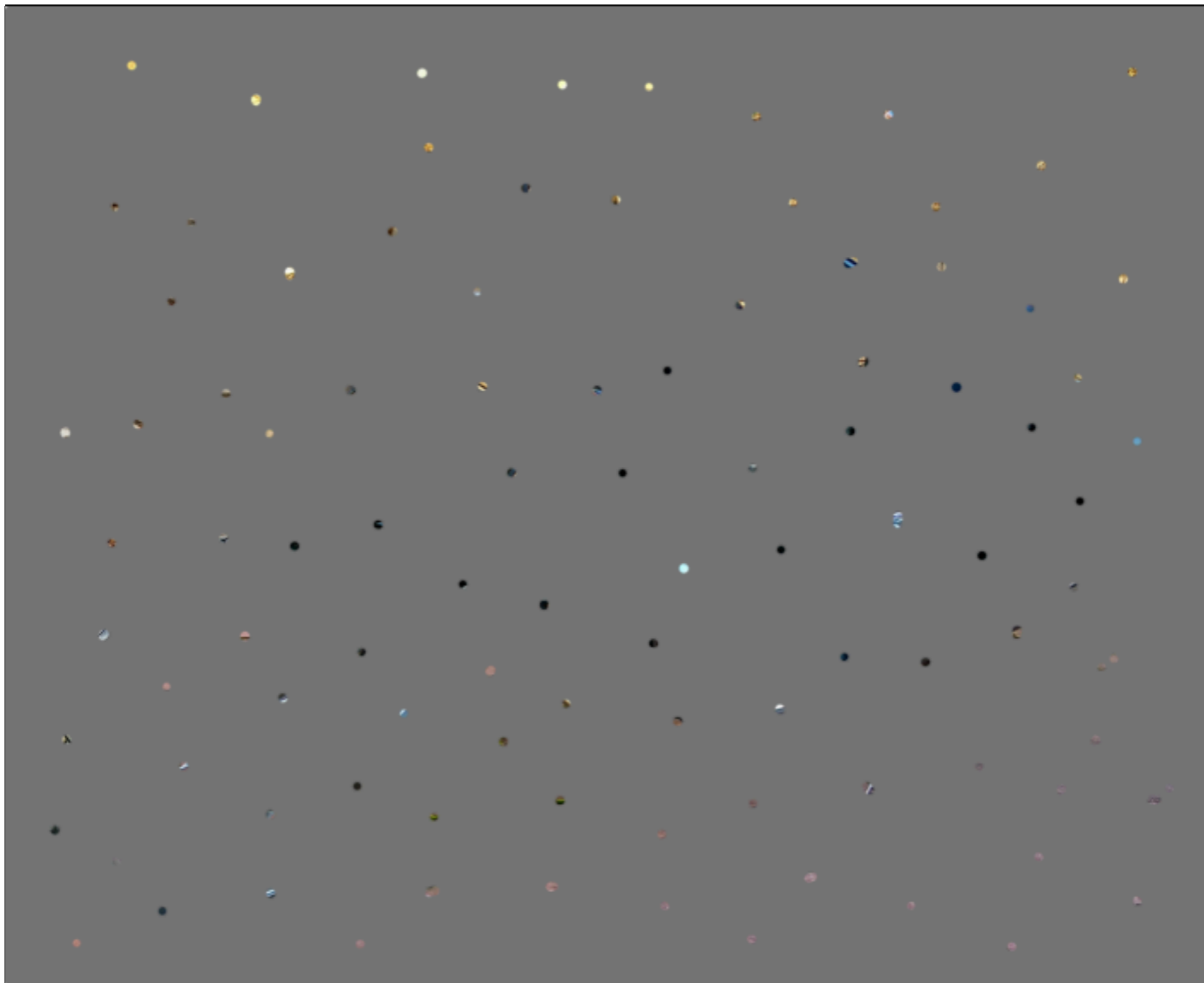
Kettlebell

# Úloha statistické indukce



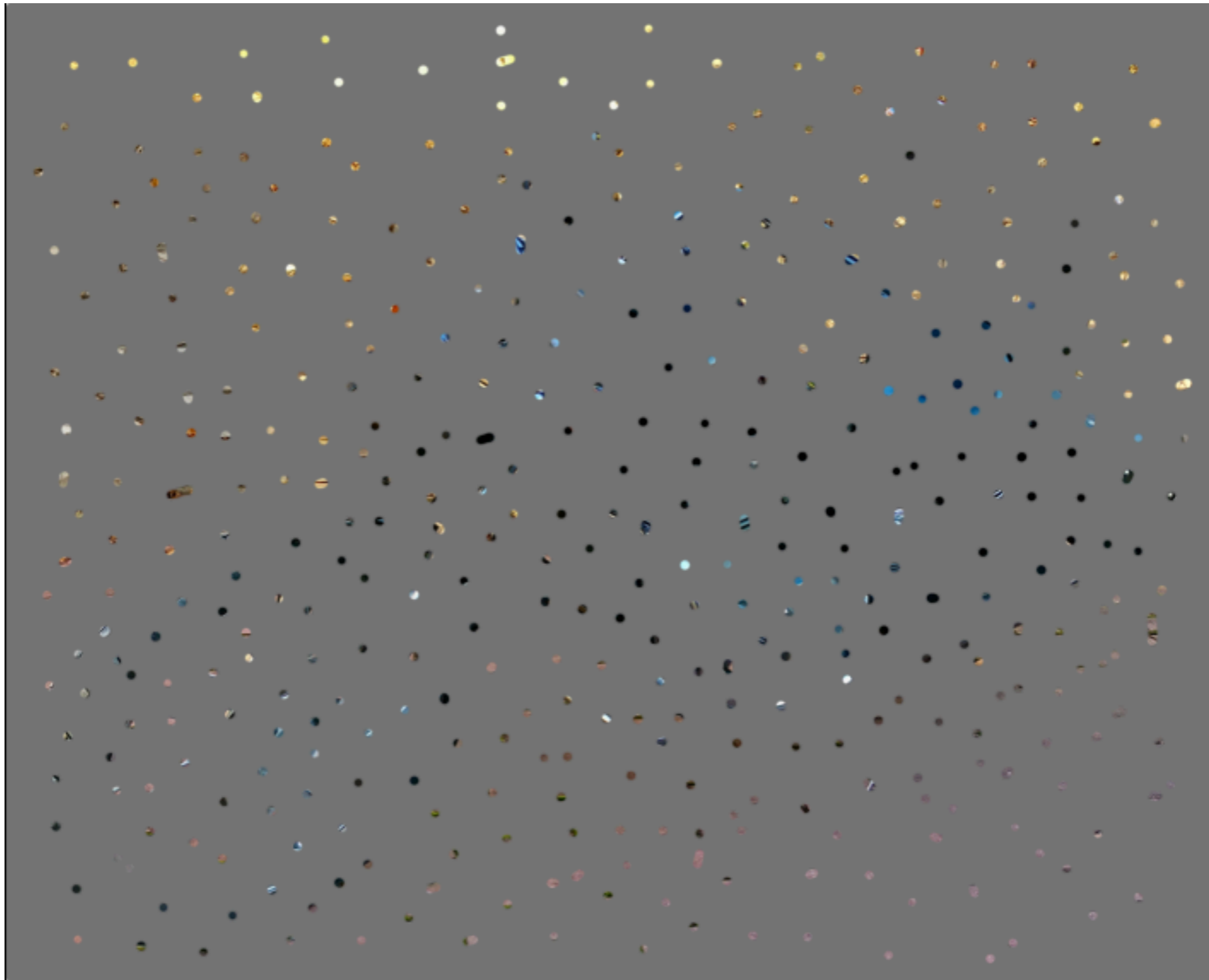
$n=50$

# Úloha statistické indukce



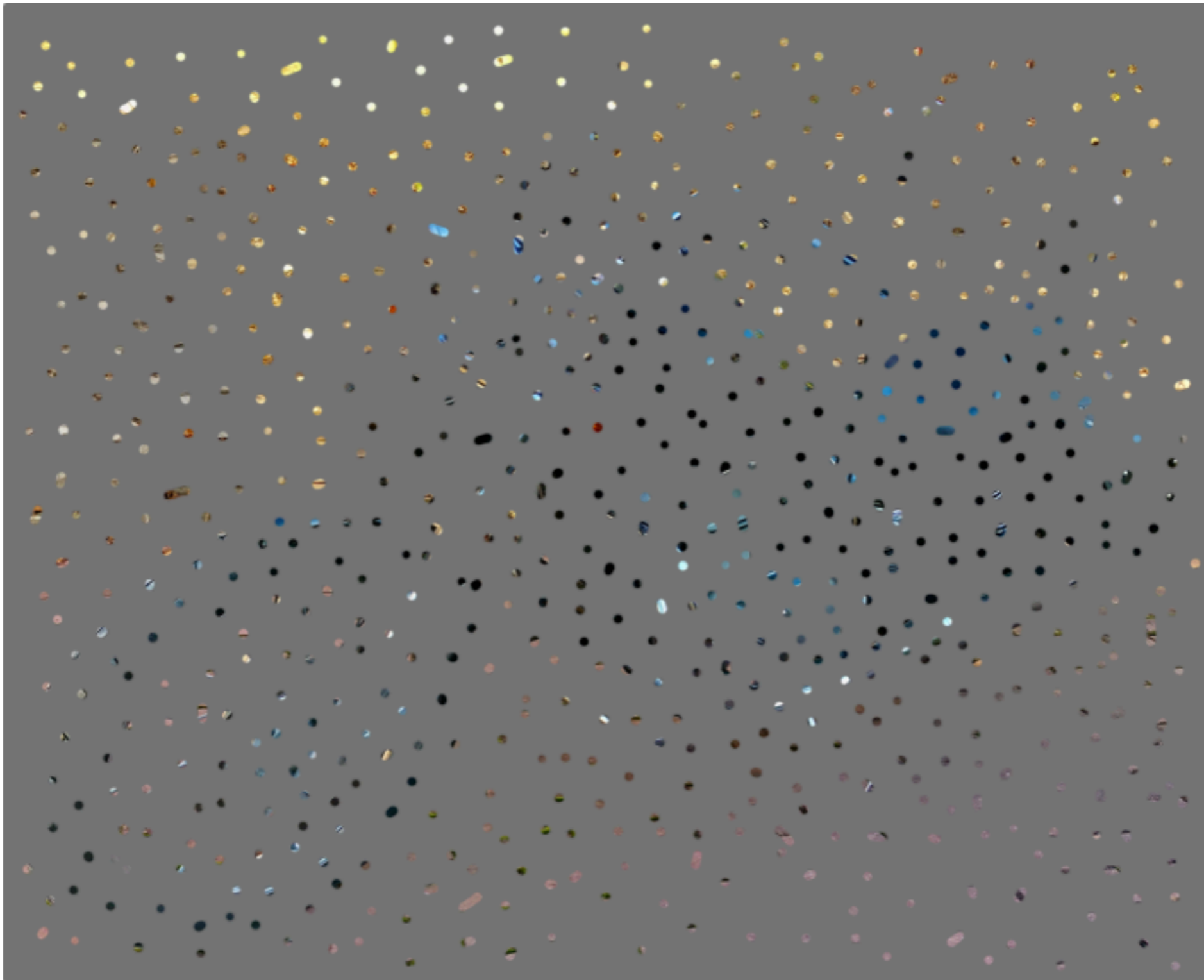
$n=100$

# Úloha statistické indukce



$n=300$

# Úloha statistické indukce



$n=500$



# Úloha statistické indukce



Kawasaki  
EN 500



# Úloha statistické indukce

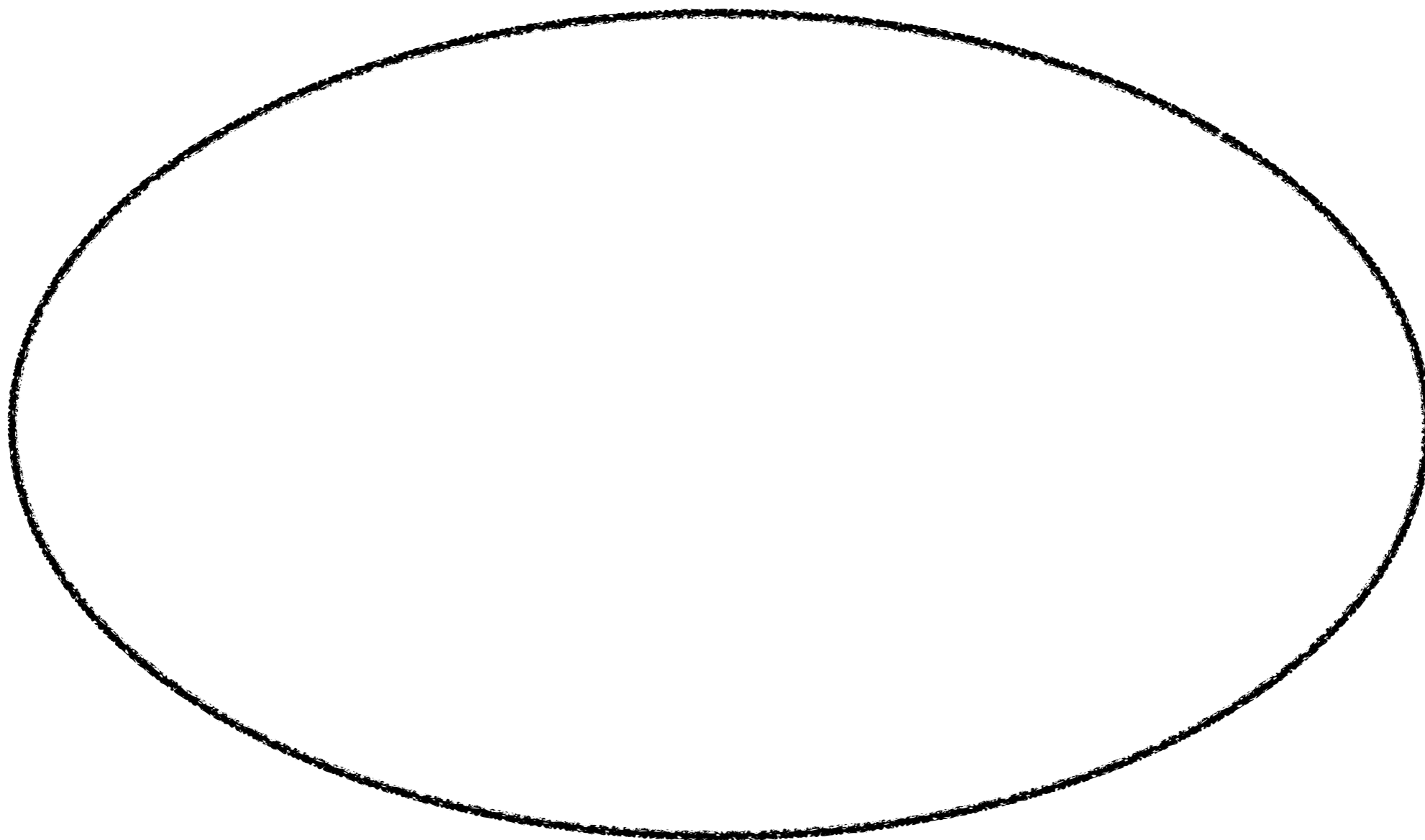


Kawasaki  
EN 500



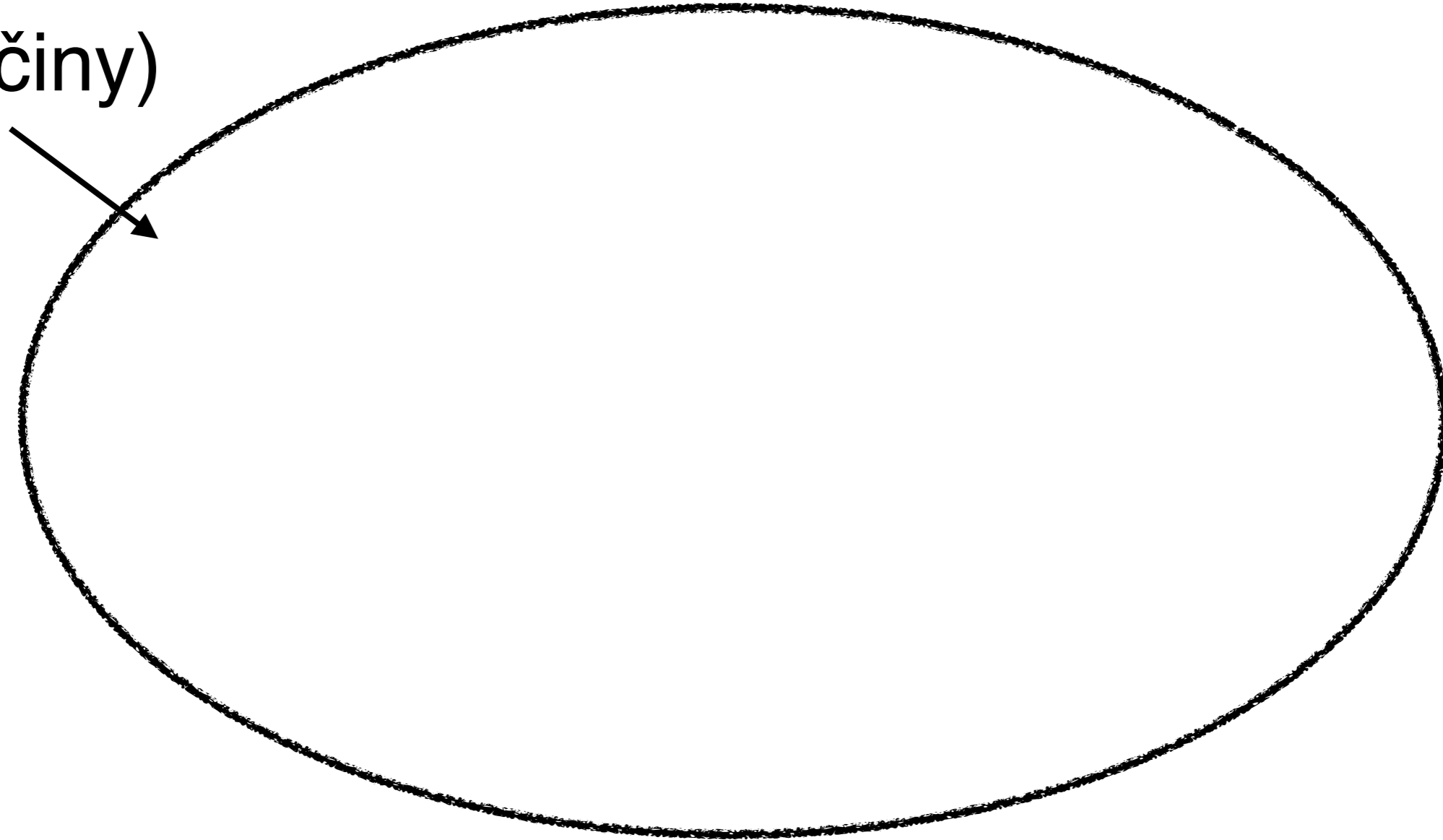


# Úloha statistické indukce



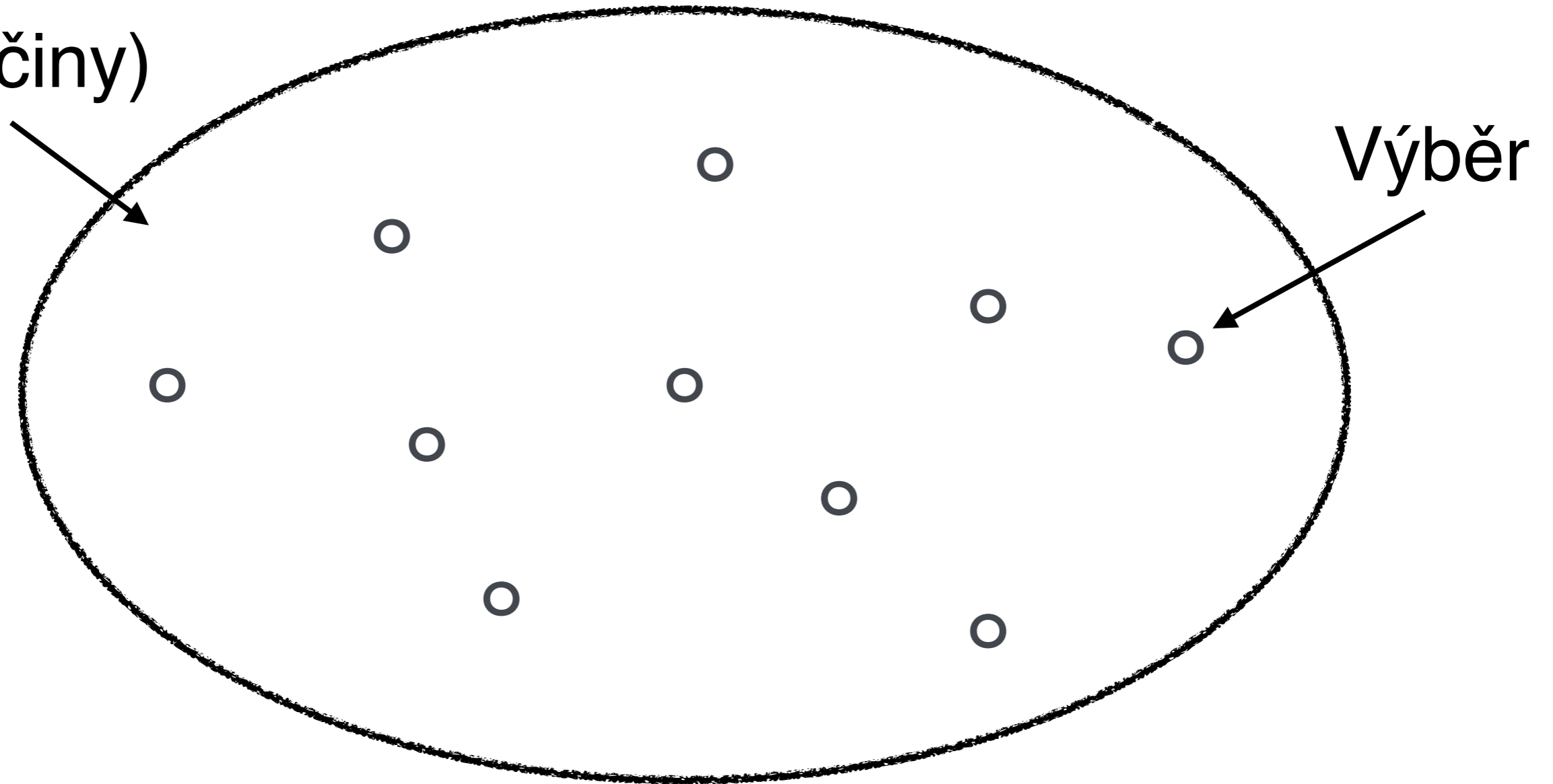
# Úloha statistické indukce

Základní soubor - nositel sledovaného znaku  
(veličiny)



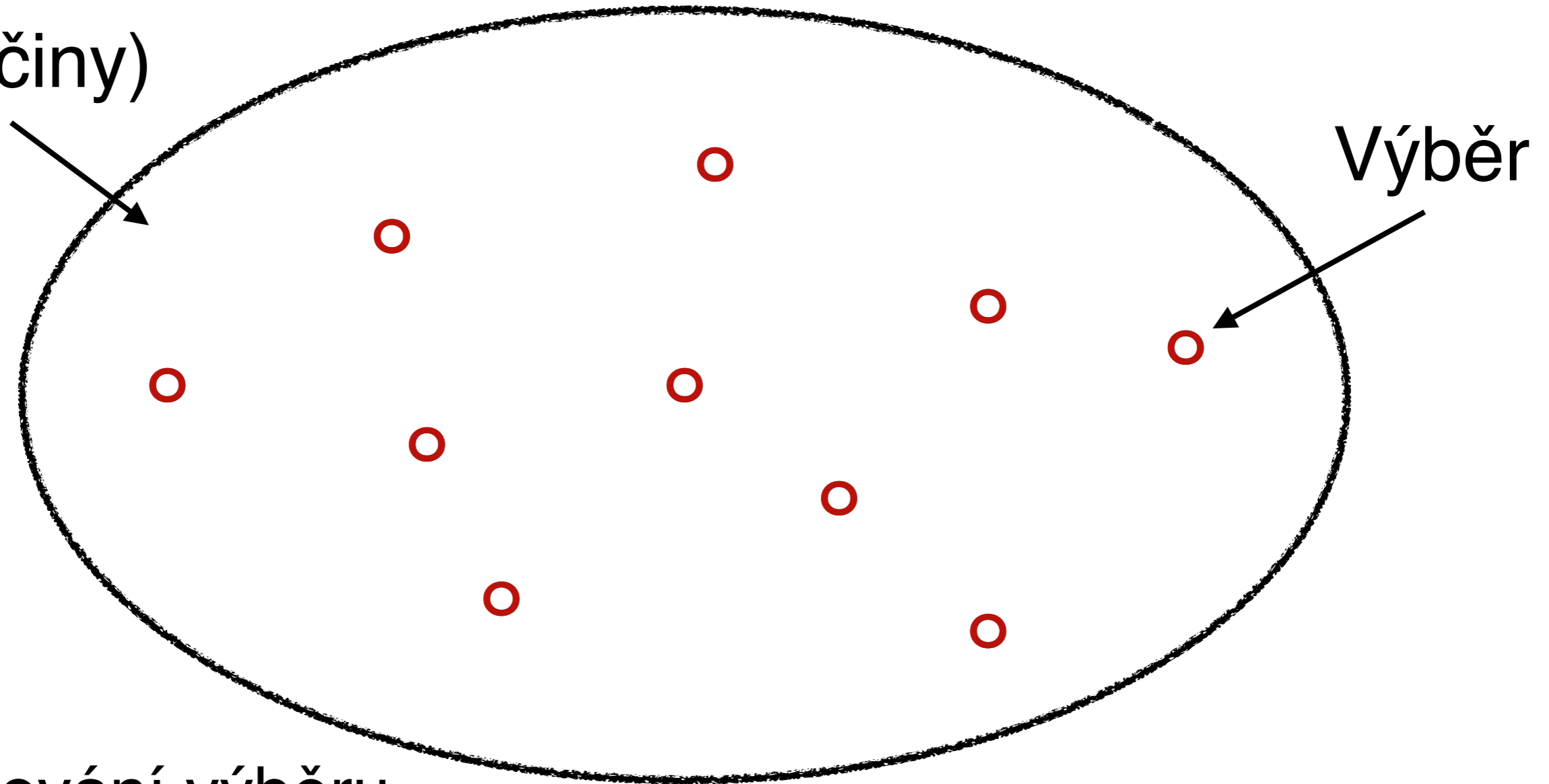
# Úloha statistické indukce

Základní soubor - nositel sledovaného znaku  
(veličiny)



# Úloha statistické indukce

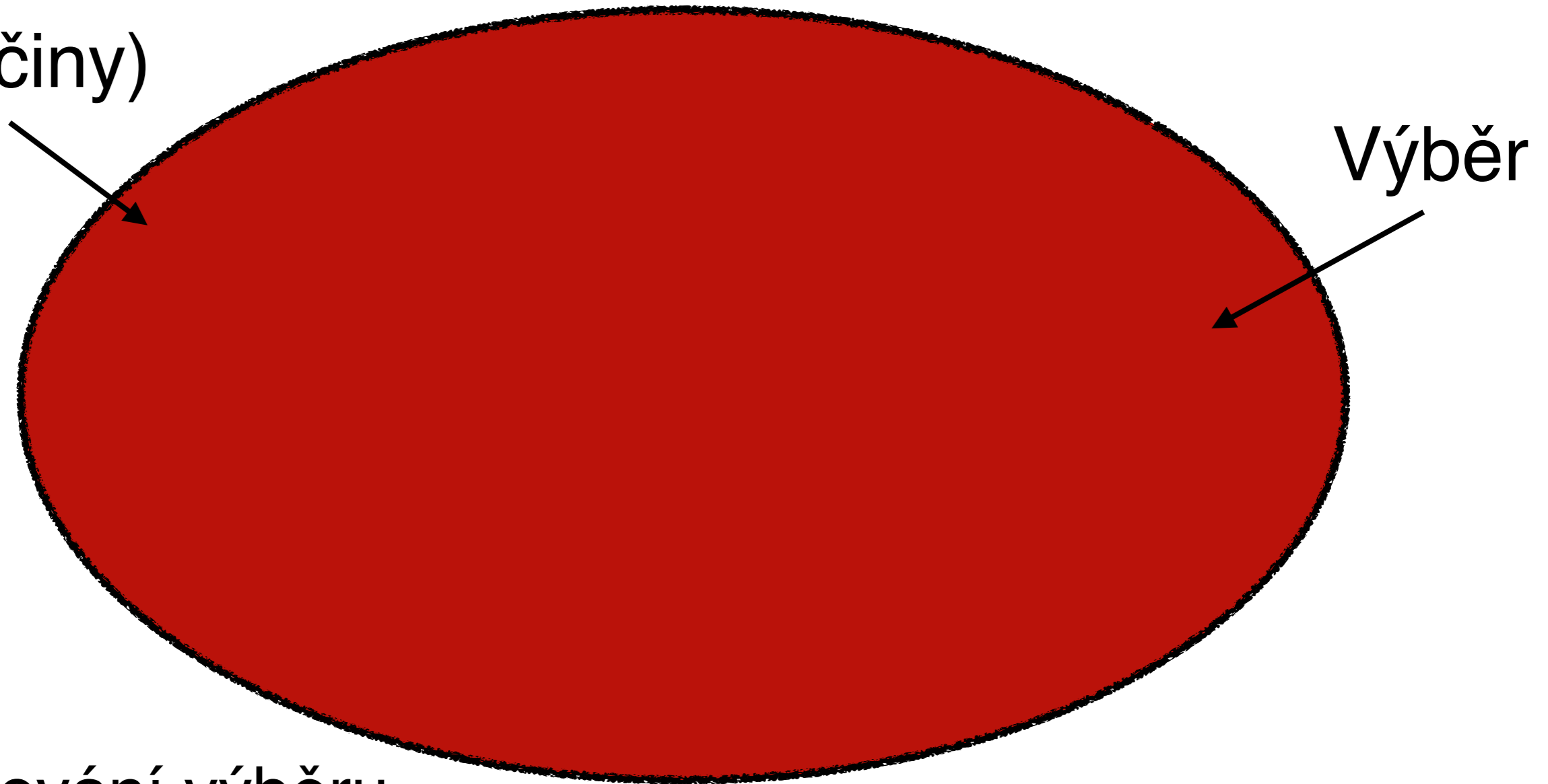
Základní soubor - nositel sledovaného znaku  
(veličiny)



Pozorování výběru  
(měření sledovaného znaku)  $\Rightarrow$  Zjištění vlastností výběru

# Úloha statistické indukce

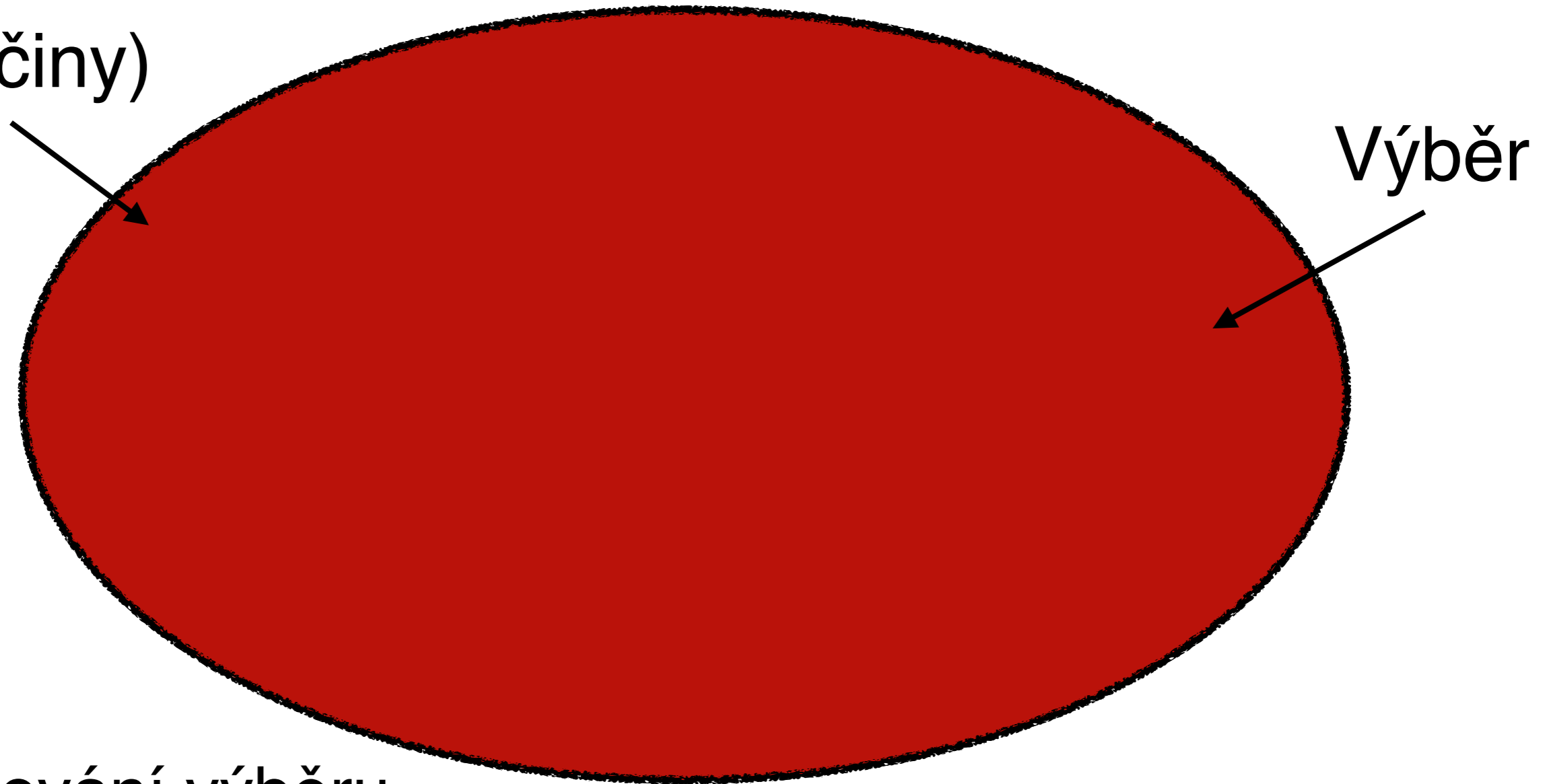
Základní soubor - nositel sledovaného znaku  
(veličiny)



Pozorování výběru  
(měření sledovaného znaku)  $\Rightarrow$  Zjištění vlastností výběru  
 $\Rightarrow$  zobecnění na celý základní soubor

# Úloha statistické indukce

Základní soubor - nositel sledovaného znaku  
(veličiny)



Pozorování výběru

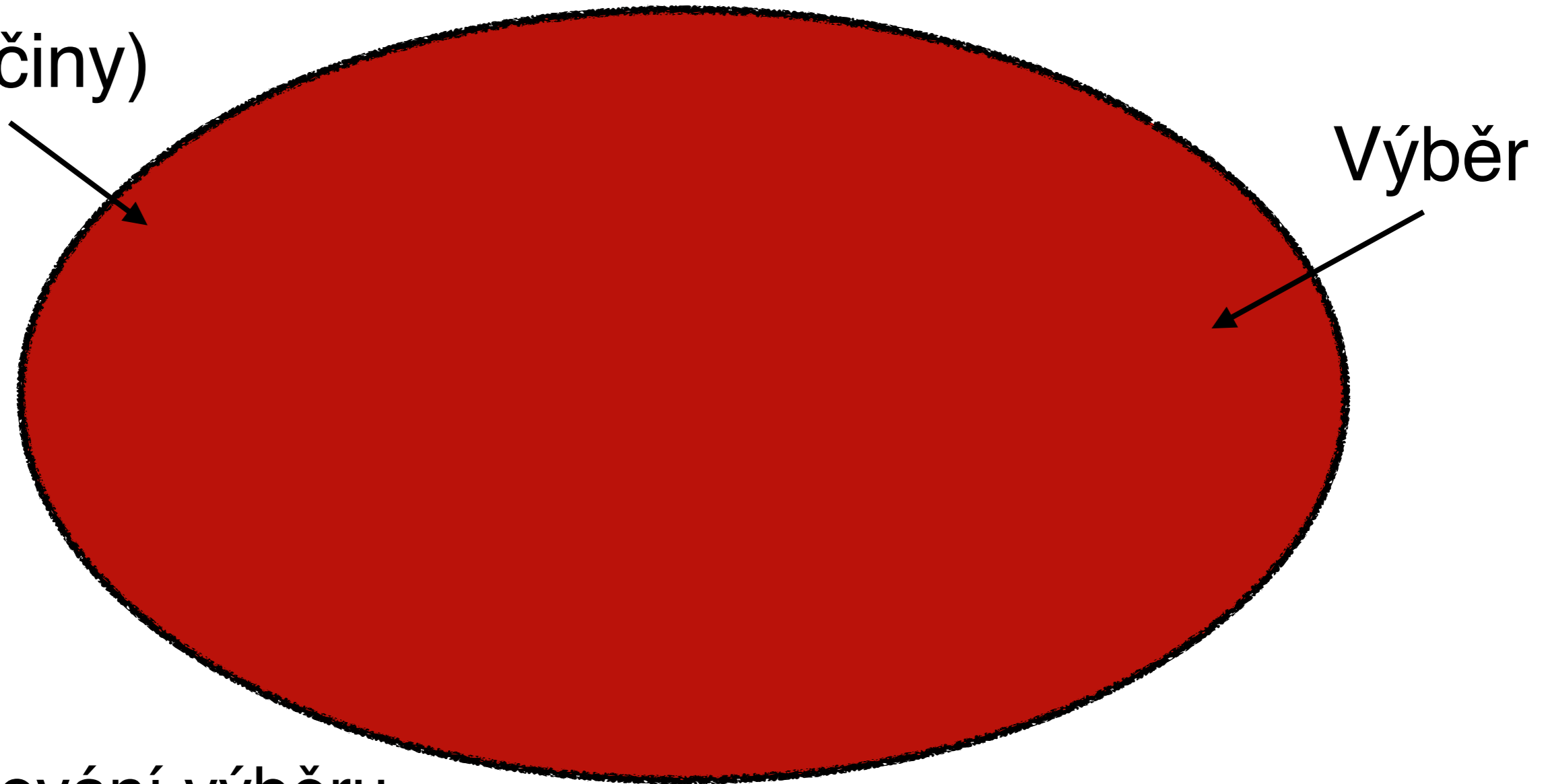
(měření sledovaného znaku)  $\Rightarrow$  Zjištění vlastností výběru

$\Rightarrow$  zobecnění na celý základní soubor



# Úloha statistické indukce

Základní soubor - nositel sledovaného znaku  
(veličiny)

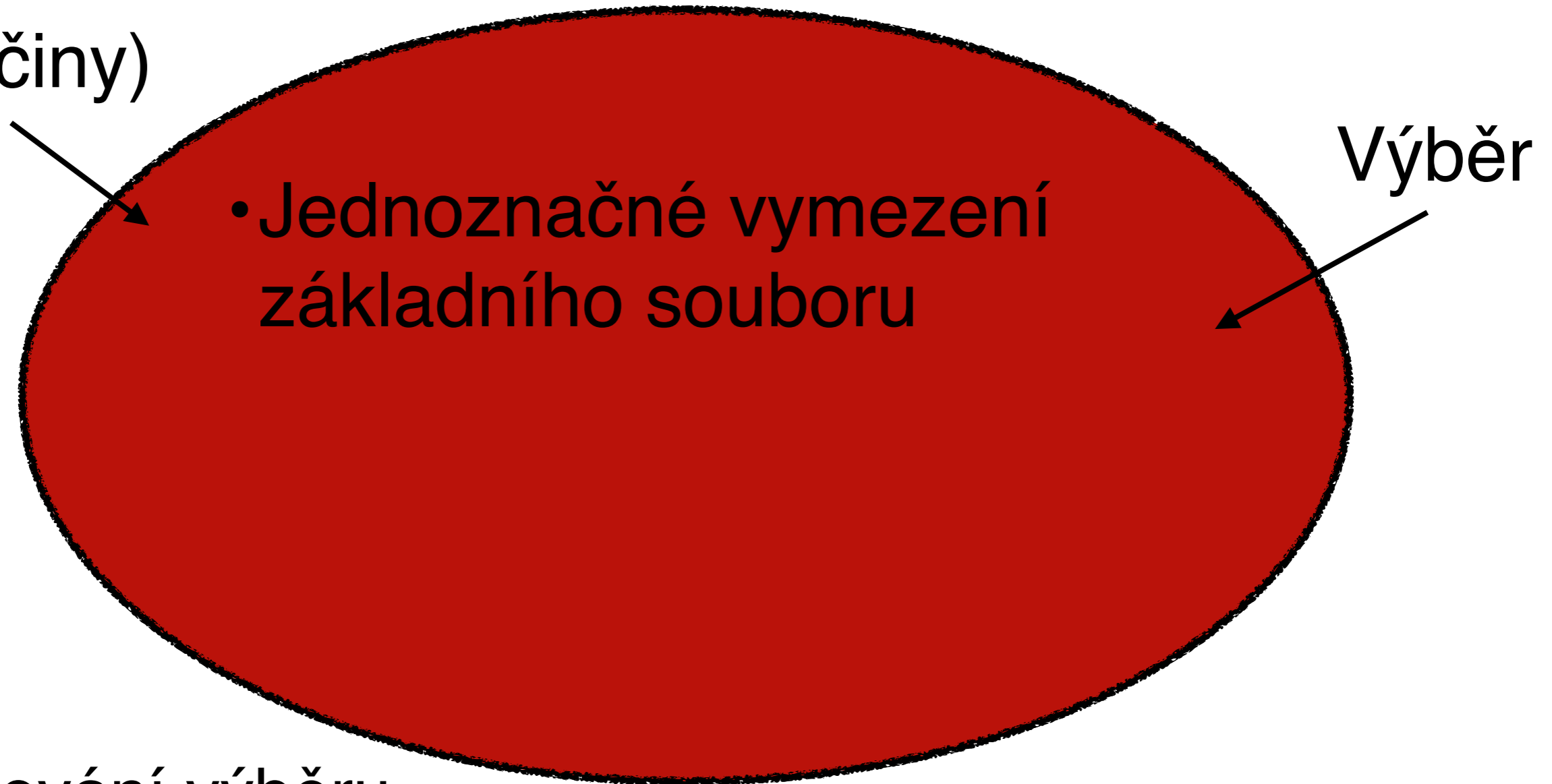


Pozorování výběru  
(měření sledovaného znaku)  $\Rightarrow$  Zjištění vlastností výběru  
 $\Rightarrow$  zobecnění na celý základní soubor



# Úloha statistické indukce

Základní soubor - nositel sledovaného znaku  
(veličiny)



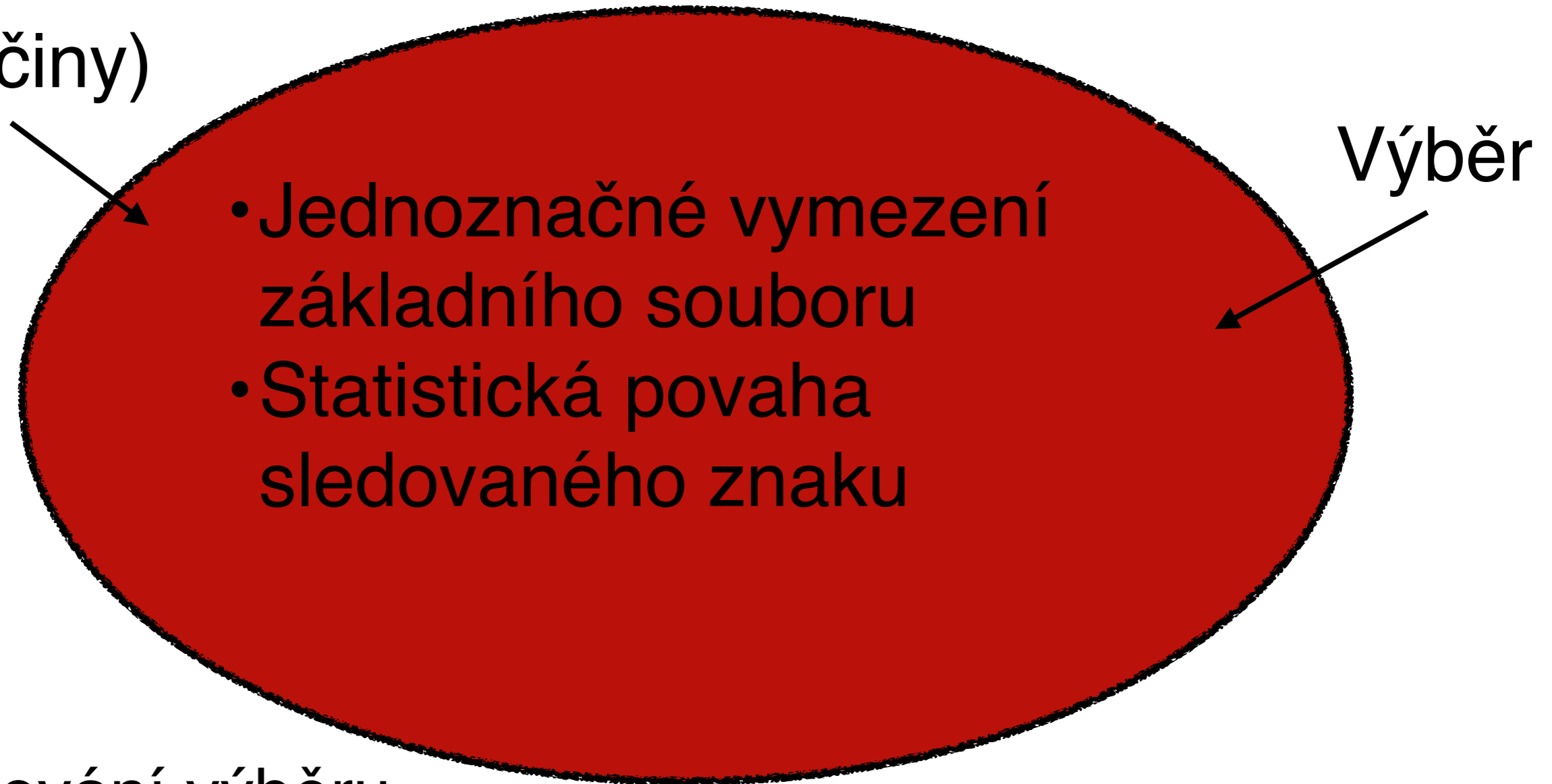
Pozorování výběru

(měření sledovaného znaku) => Zjištění vlastností výběru

=> zobecnění na celý základní soubor

# Úloha statistické indukce

Základní soubor - nositel sledovaného znaku  
(veličiny)



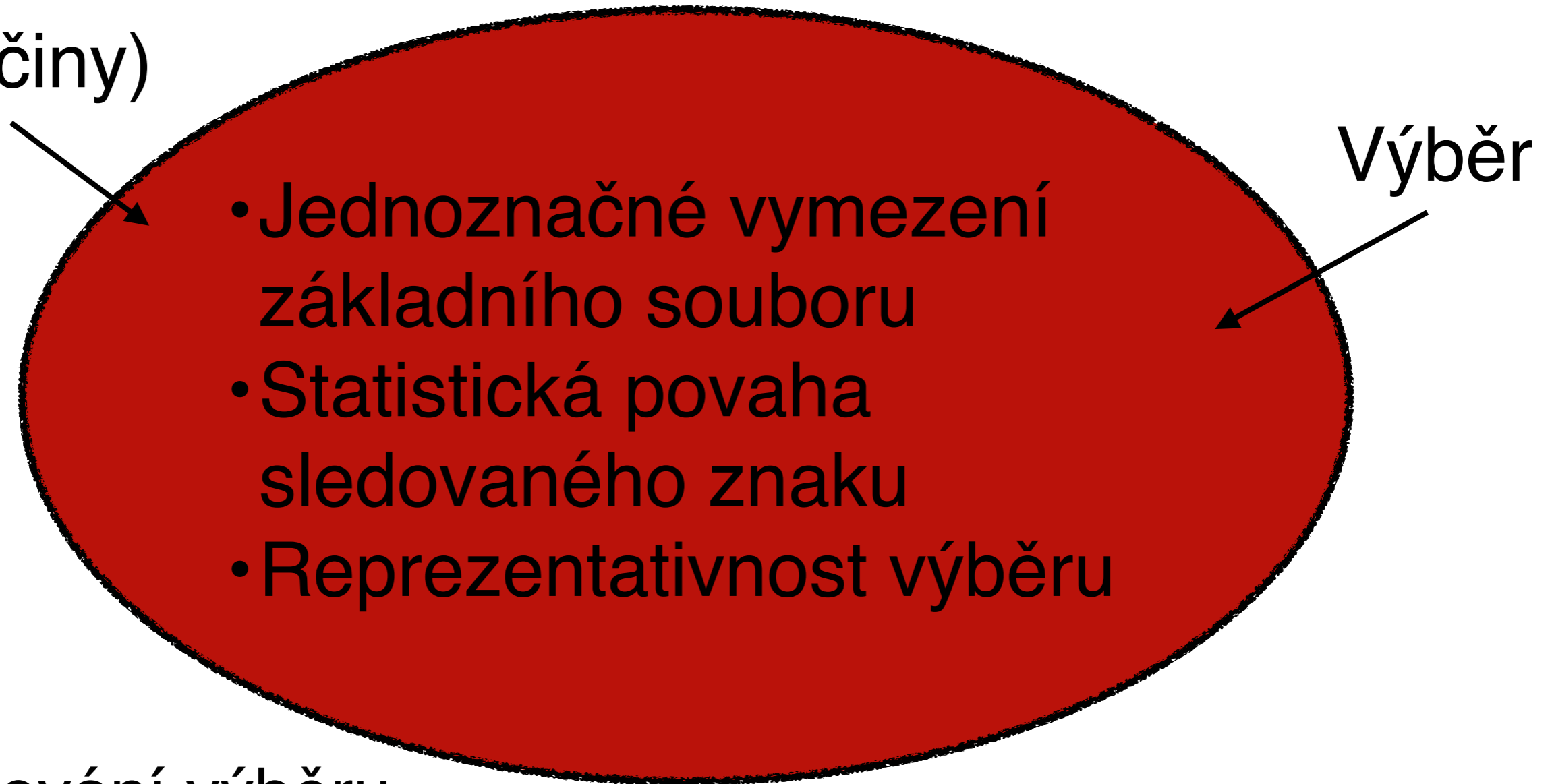
Pozorování výběru

(měření sledovaného znaku) => Zjištění vlastností výběru

=> zobecnění na celý základní soubor

# Úloha statistické indukce

Základní soubor - nositel sledovaného znaku  
(veličiny)



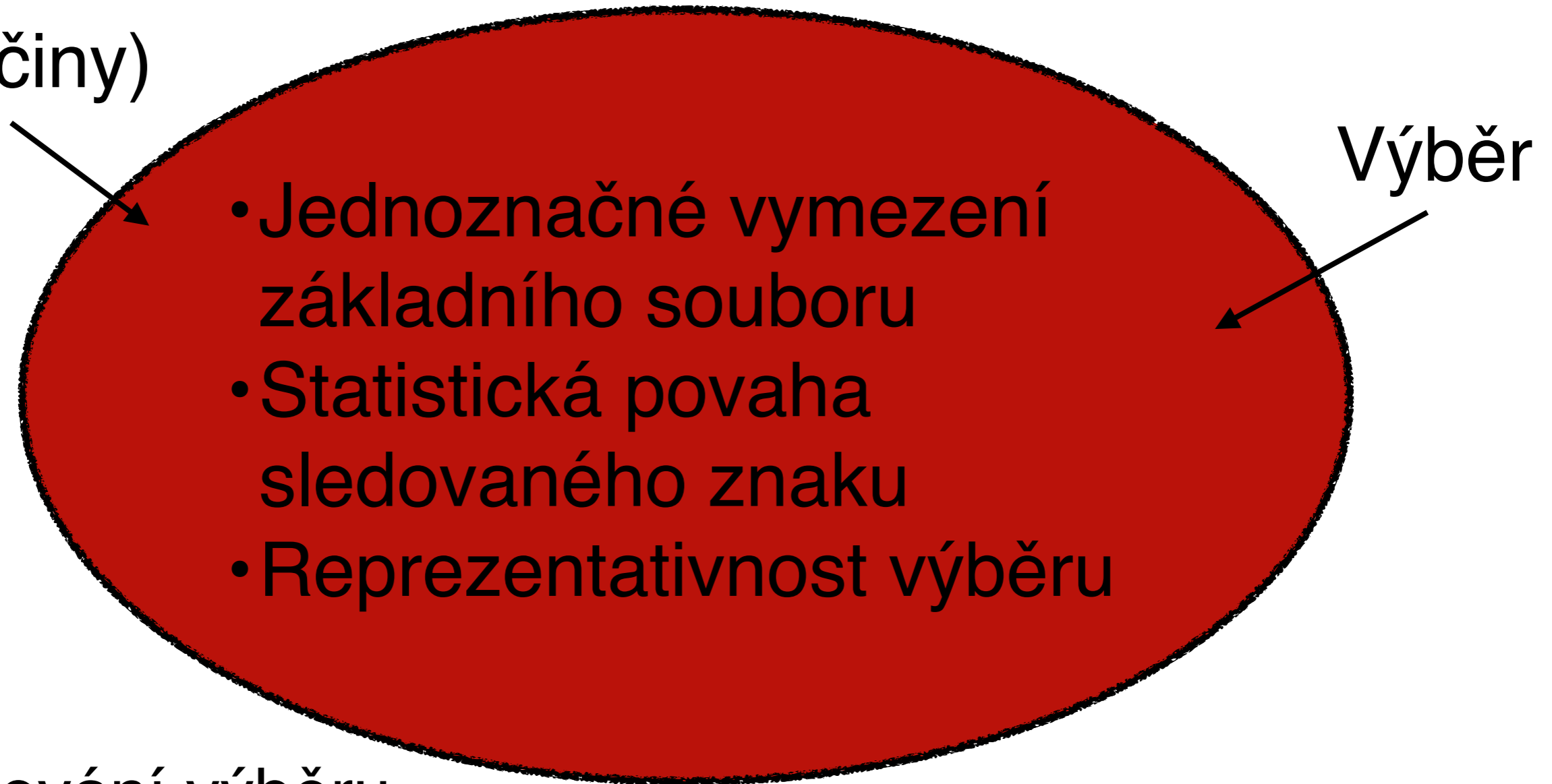
Pozorování výběru

(měření sledovaného znaku) => Zjištění vlastností výběru

=> zobecnění na celý základní soubor

# Úloha statistické indukce

Základní soubor - nositel sledovaného znaku  
(veličiny)



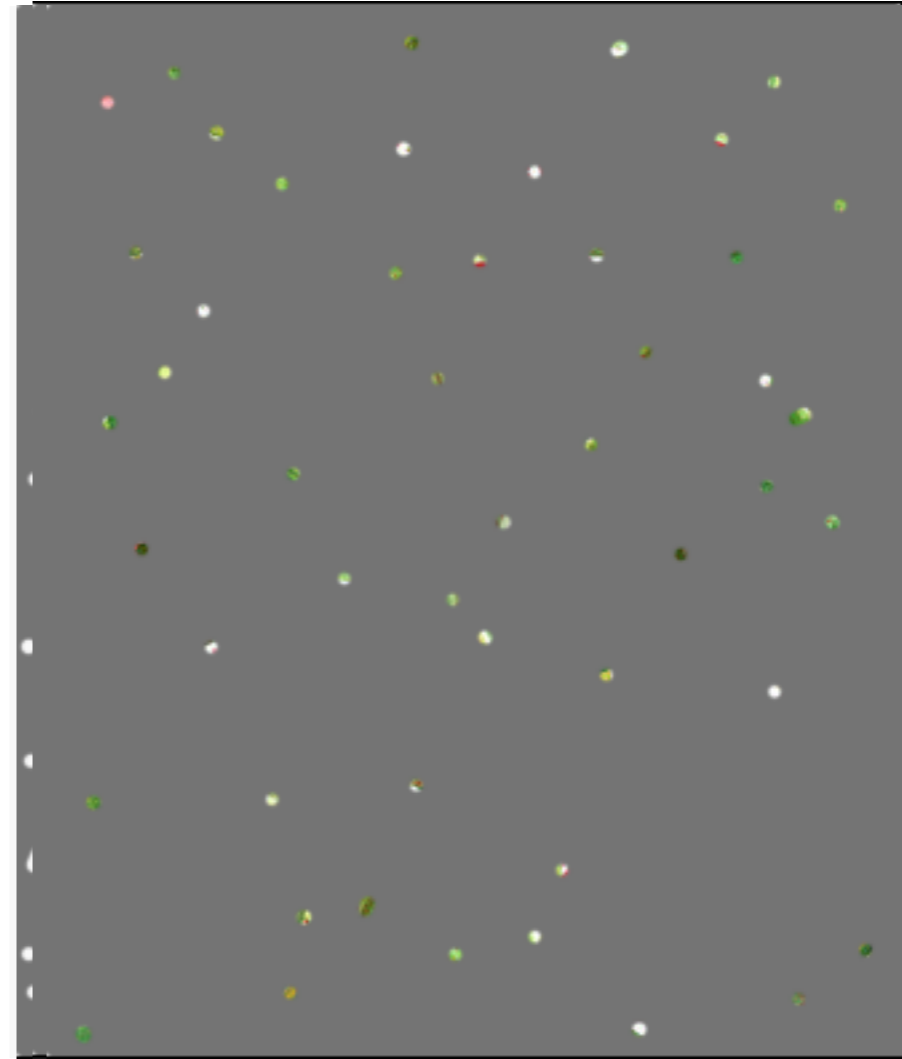
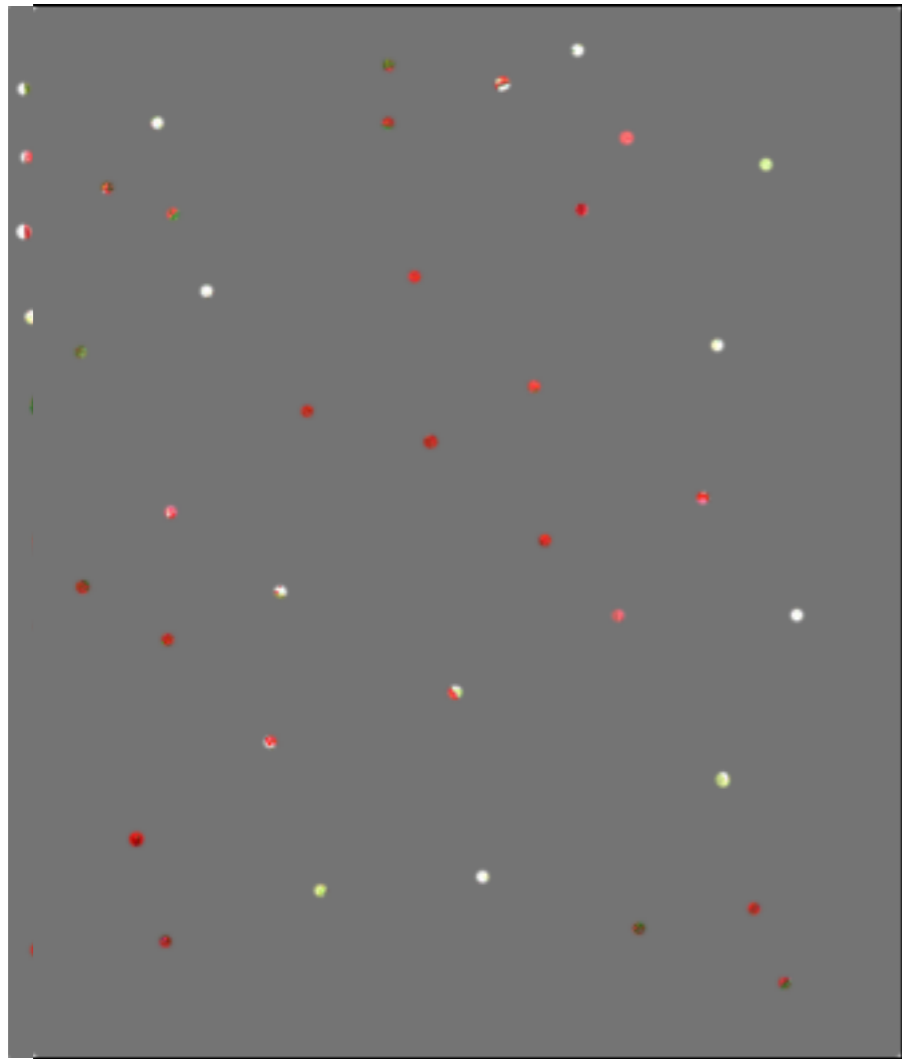
Pozorování výběru

(měření sledovaného znaku) => Zjištění vlastností výběru

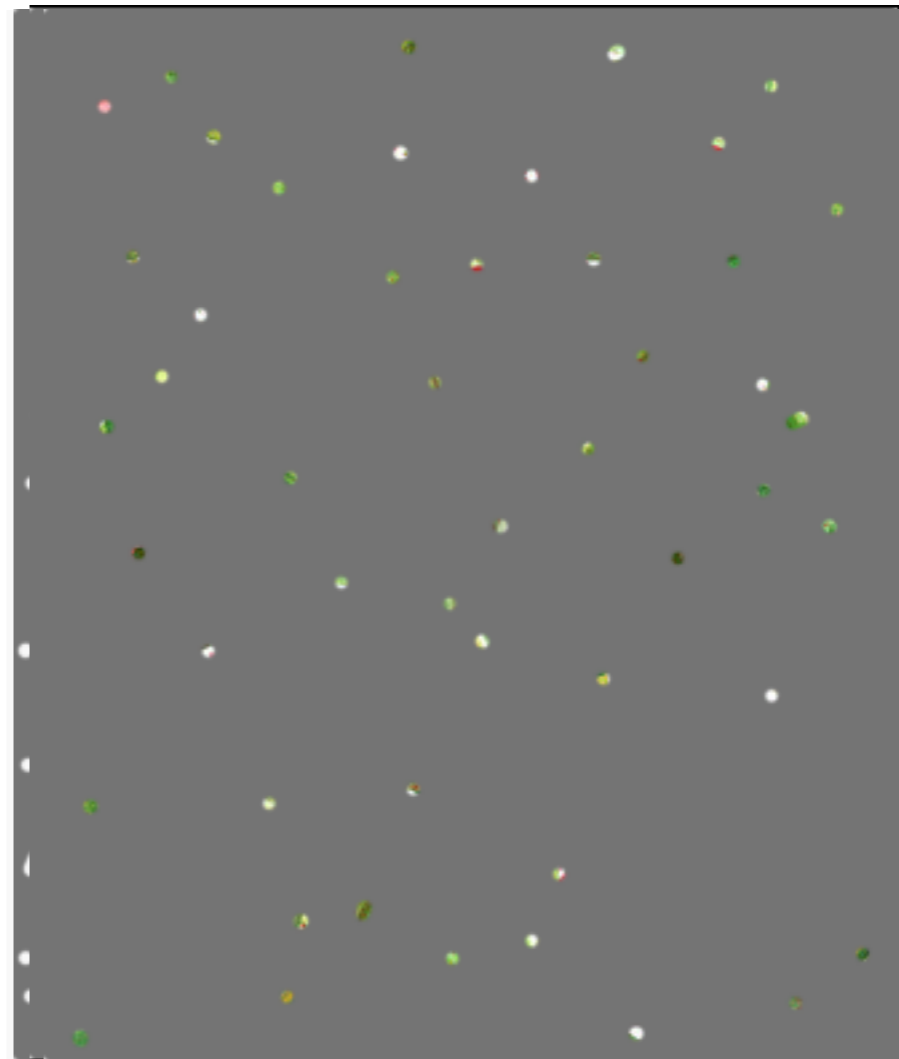
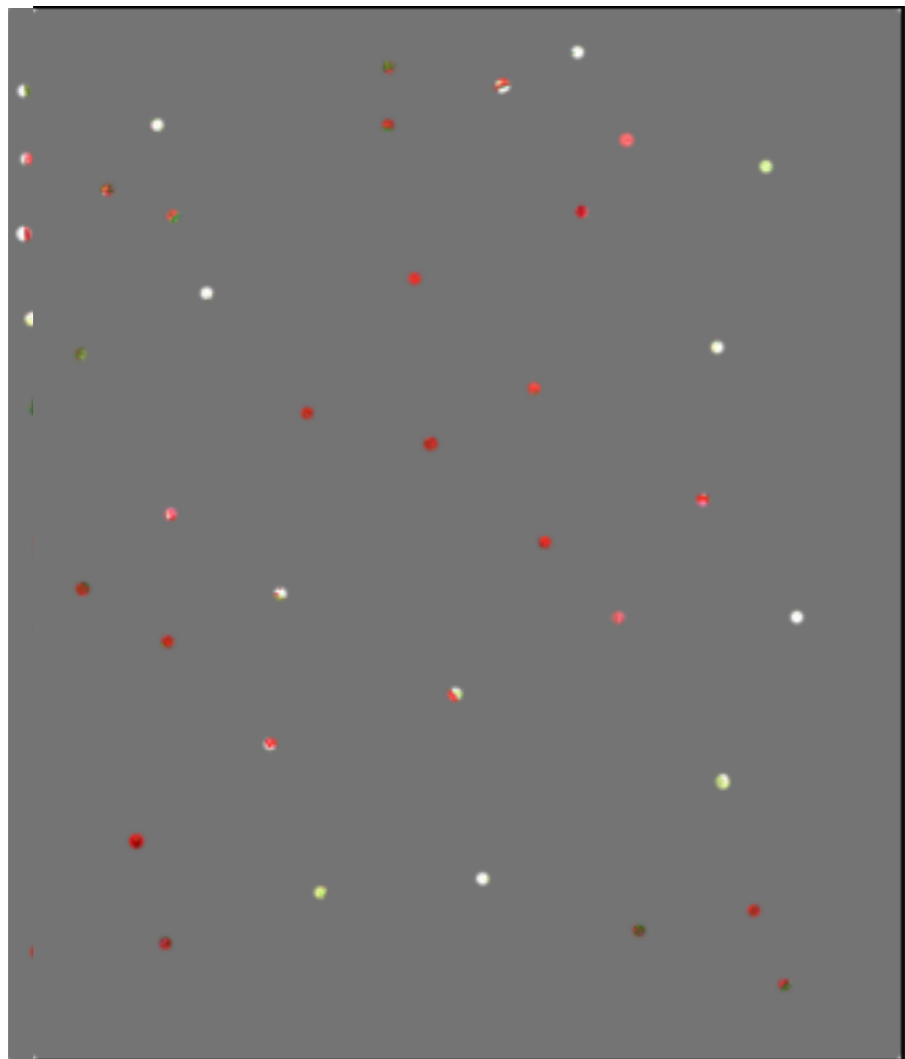
=> zobecnění na celý základní soubor



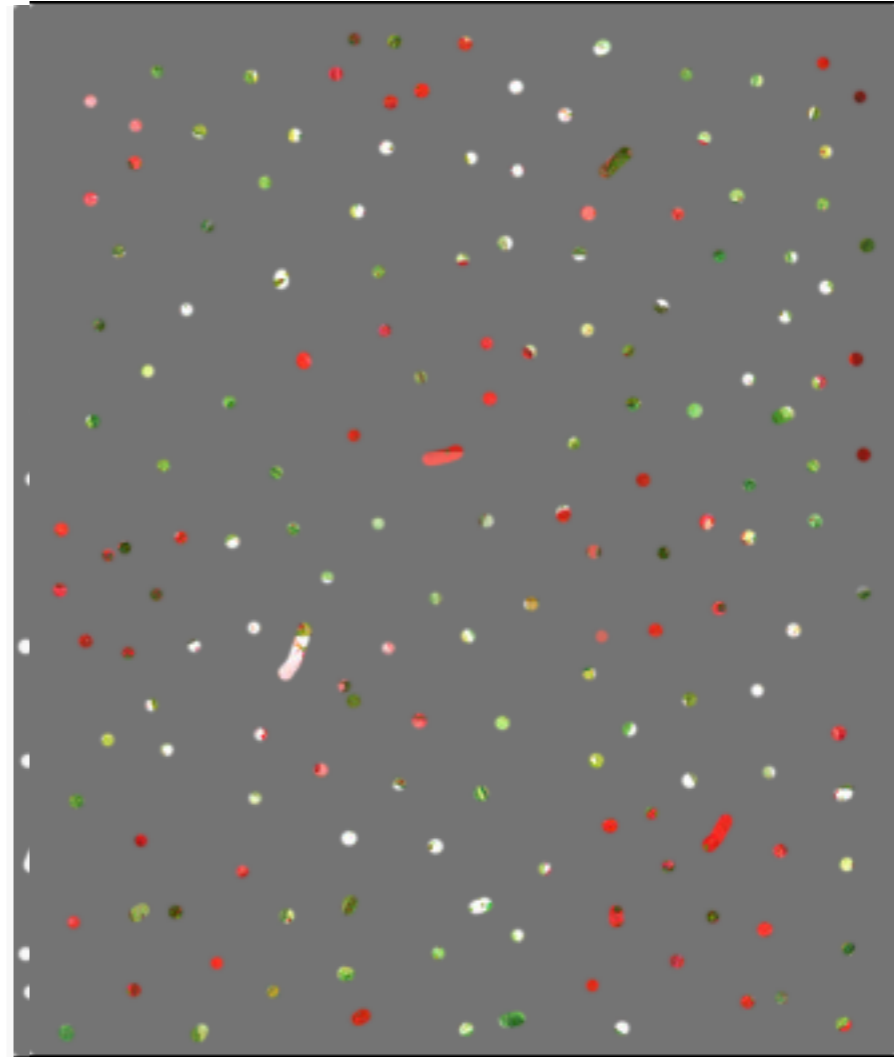
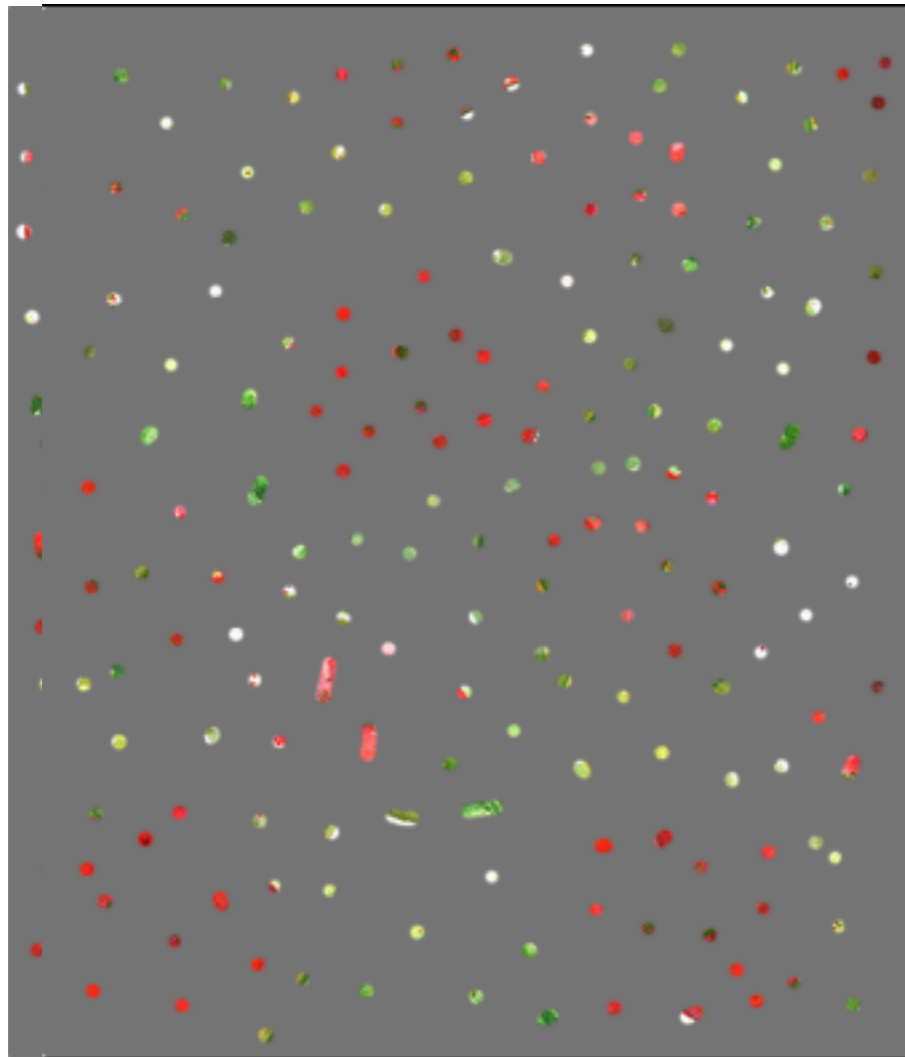
# Úloha statistické indukce



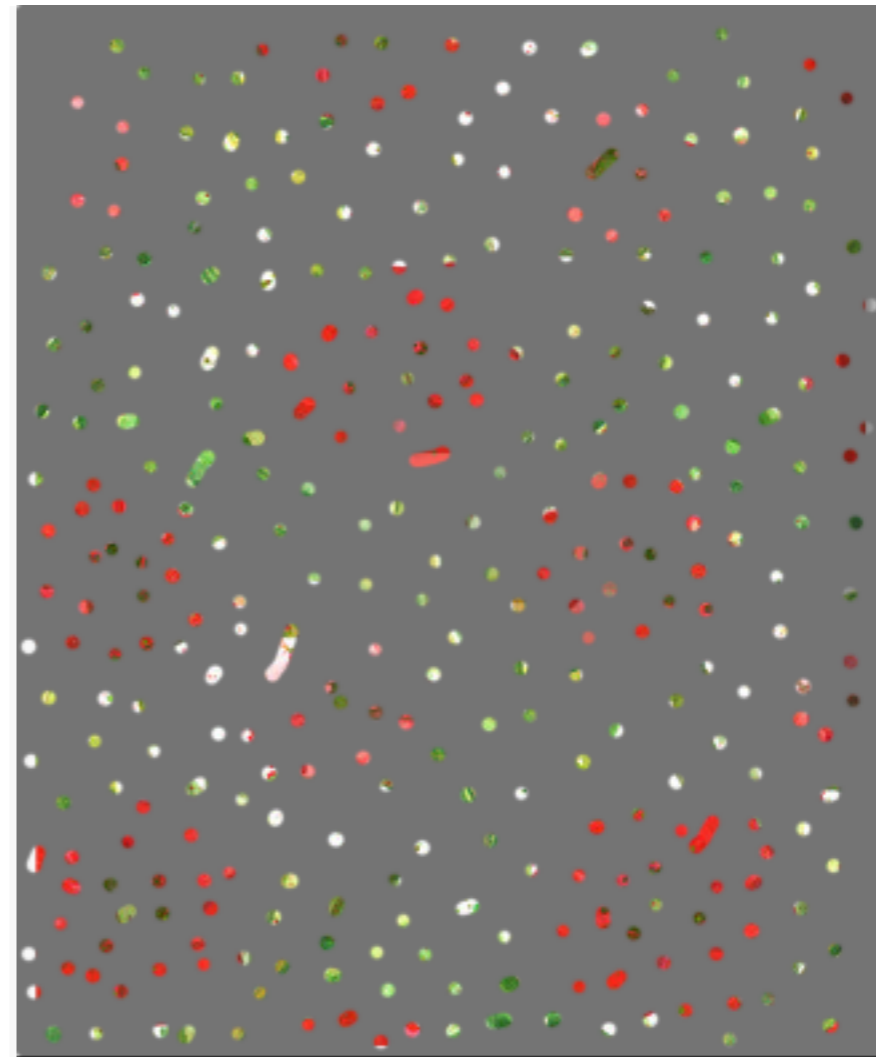
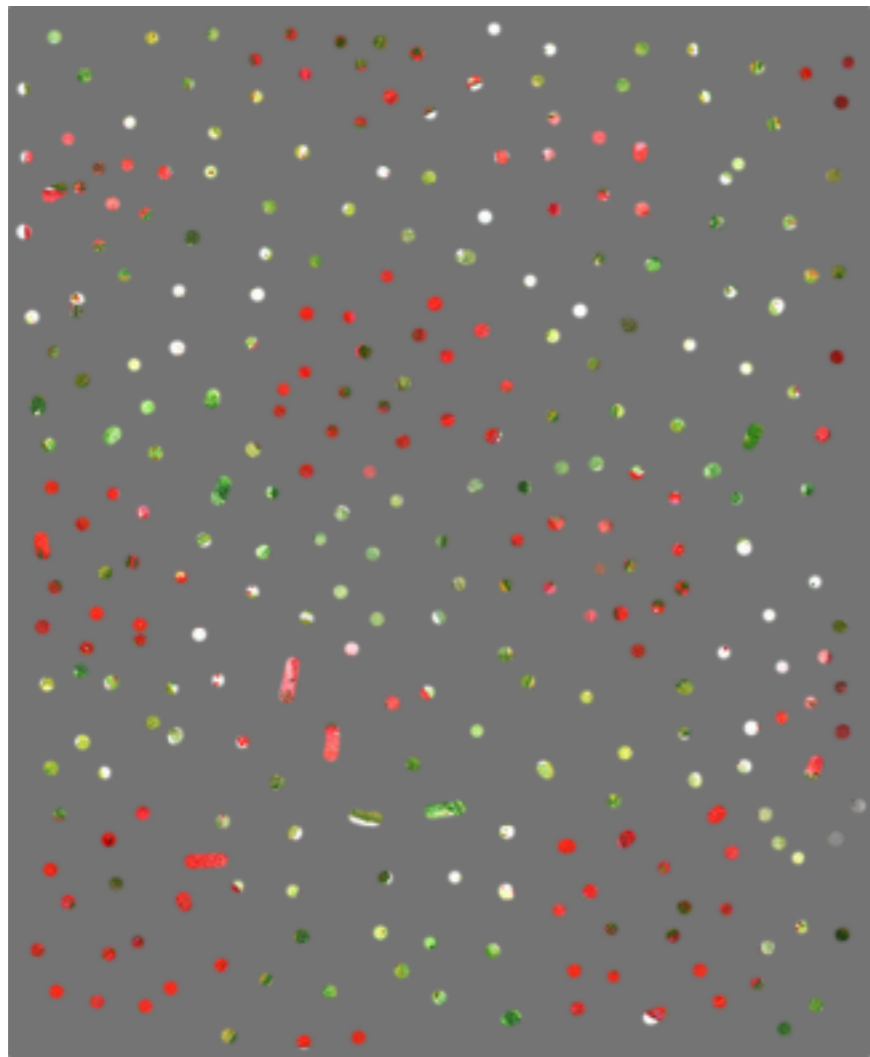
# Úloha statistické indukce



# Úloha statistické indukce



# Úloha statistické indukce





# Úloha statistické indukce



# Statistické charakteristiky

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .

| Pravděpodobnostní charakteristiky                             | Výběrové charakteristiky   |
|---|--|
| Střední hodnota<br>$E(X) = \int_{-\infty}^{\infty} x f(x) dx$ | Výběrový průměr<br>$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$              |
| Momenty<br>$\mu_k(X) = E(X - E(X))^k$                         | Výběrové momenty<br>$m_k(X) = \sum_{i=1}^n (x_i - \bar{X})^k$            |
| Rozptyl<br>$var(X) = E(X - E(X))^2$                           | Výběrový rozptyl<br>$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2$ |
| Kvantily<br>$\tilde{x}_{100\alpha}$                           | Výběrové kvantily<br>$X_{([np]+1)}$                                      |

# Statistické charakteristiky

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .

| Pravděpodobnostní charakteristiky                             | Výběrové charakteristiky   |
|---|--|
| Střední hodnota<br>$E(X) = \int_{-\infty}^{\infty} x f(x) dx$ | Výběrový průměr<br>$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$              |
| Momenty<br>$\mu_k(X) = E(X - E(X))^k$                         | Výběrové momenty<br>$m_k(X) = \sum_{i=1}^n (x_i - \bar{X})^k$            |
| Rozptyl<br>$var(X) = E(X - E(X))^2$                           | Výběrový rozptyl<br>$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2$ |
| Kvantily<br>$\tilde{x}_{100\alpha}$                           | Výběrové kvantily<br>$X_{([np]+1)}$                                      |



# Statistické charakteristiky

Za předpokladu, že náhodný výběr je nezávislý a je z normálního rozdělení, t.j.  $X_1, X_2, \dots, X_n$  jsou i.i.d. a  $X_k \sim N(\mu, \sigma^2)$ ,  $k = 1, 2, \dots, n$ , lze určit rozdělení pravděpodobnosti některých charakteristik:

# Statistické charakteristiky

Za předpokladu, že náhodný výběr je nezávislý a je z normálního rozdělení, t.j.  $X_1, X_2, \dots, X_n$  jsou i.i.d. a  $X_k \sim N(\mu, \sigma^2)$ ,  $k = 1, 2, \dots, n$ , lze určit rozdělení pravděpodobnosti některých charakteristik:

- Pokud je  $\mu$  a  $\sigma^2$  známé, má výběrový průměr  $\bar{X}_n$  rozdělení  $N(\mu, \sigma^2/n)$

# Statistické charakteristiky

Za předpokladu, že náhodný výběr je nezávislý a je z normálního rozdělení, t.j.  $X_1, X_2, \dots, X_n$  jsou i.i.d. a  $X_k \sim N(\mu, \sigma^2)$ ,  $k = 1, 2, \dots, n$ , lze určit rozdělení pravděpodobnosti některých charakteristik:

- Pokud je  $\mu$  a  $\sigma^2$  známé, má výběrový průměr  $\bar{X}_n$  rozdělení  $N(\mu, \sigma^2/n)$
- Pokud  $\mu$  a  $\sigma^2$  neznáme, má veličina  $T = (X - \bar{X})/s$  tzv. Studentovo neboli  $t$ -rozdělení  $t(n-1)$

# Statistické charakteristiky

Za předpokladu, že náhodný výběr je nezávislý a je z normálního rozdělení, t.j.  $X_1, X_2, \dots, X_n$  jsou i.i.d. a  $X_k \sim N(\mu, \sigma^2)$ ,  $k = 1, 2, \dots, n$ , lze určit rozdělení pravděpodobnosti některých charakteristik:

- Pokud je  $\mu$  a  $\sigma^2$  známé, má výběrový průměr  $\bar{X}_n$  rozdělení  $N(\mu, \sigma^2/n)$
- Pokud  $\mu$  a  $\sigma^2$  neznáme, má veličina  $T = (X - \bar{X})/s$  tzv. Studentovo neboli  $t$ -rozdělení  $t(n-1)$
- Veličina  $S^2 = (n-1) \cdot s^2 / \sigma^2$  má  $\chi^2(n-1)$  rozdělení (o  $n-1$  stupních volnosti)

# Statistické charakteristiky

Za předpokladu, že náhodný výběr je nezávislý a je z normálního rozdělení, t.j.  $X_1, X_2, \dots, X_n$  jsou i.i.d. a  $X_k \sim N(\mu, \sigma^2)$ ,  $k = 1, 2, \dots, n$ , lze určit rozdělení pravděpodobnosti některých charakteristik:

- Pokud je  $\mu$  a  $\sigma^2$  známé, má výběrový průměr  $\bar{X}_n$  rozdělení  $N(\mu, \sigma^2/n)$
- Pokud  $\mu$  a  $\sigma^2$  neznáme, má veličina  $T = (X - \bar{X})/s$  tzv. Studentovo neboli  $t$ -rozdělení  $t(n-1)$
- Veličina  $S^2 = (n-1) \cdot s^2 / \sigma^2$  má  $\chi^2(n-1)$  rozdělení (o  $n-1$  stupních volnosti)





# Statistické charakteristiky

Další důležité výběrové charakteristiky:

- Výběrová šikmost (skewness):  $Skew(X) = \frac{m_3(X)}{m_2^{3/2}(X)}$

# Statistické charakteristiky

Další důležité výběrové charakteristiky:

- Výběrová šikmost (skewness):  $Skew(X) = \frac{m_3(X)}{m_2^{3/2}(X)}$   
pro  $X \sim N(\mu, \sigma^2)$  je  
 $E(Skew(X)) = 0$        $var(Skew(X)) = \frac{6(n-2)}{(n+1)(n+3)}$

# Statistické charakteristiky

Další důležité výběrové charakteristiky:

- Výběrová šikmost (skewness):  $Skew(X) = \frac{m_3(X)}{m_2^{3/2}(X)}$   
pro  $X \sim N(\mu, \sigma^2)$  je  
 $E(Skew(X)) = 0$        $var(Skew(X)) = \frac{6(n-2)}{(n+1)(n+3)}$
- Výběrová špičatost (kurtosis):  $Kurt(X) = \frac{m_4(X)}{m_2^2(X)} - 3$

# Statistické charakteristiky

Další důležité výběrové charakteristiky:

- Výběrová šikmost (skewness):  $Skew(X) = \frac{m_3(X)}{m_2^{3/2}(X)}$   
pro  $X \sim N(\mu, \sigma^2)$  je

$$E(Skew(X)) = 0 \quad var(Skew(X)) = \frac{6(n-2)}{(n+1)(n+3)}$$

- Výběrová špičatost (kurtosis):  $Kurt(X) = \frac{m_4(X)}{m_2^2(X)} - 3$   
pro  $X \sim N(\mu, \sigma^2)$  je

$$E(Kurt(X)) = -\frac{6}{n+1} \quad var(Kurt(X)) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$$

# Statistické charakteristiky

Další důležité výběrové charakteristiky:

- Výběrová šikmost (skewness):  $Skew(X) = \frac{m_3(X)}{m_2^{3/2}(X)}$   
pro  $X \sim N(\mu, \sigma^2)$  je

$$E(Skew(X)) = 0 \quad var(Skew(X)) = \frac{6(n-2)}{(n+1)(n+3)}$$

- Výběrová špičatost (kurtosis):  $Kurt(X) = \frac{m_4(X)}{m_2^2(X)} - 3$   
pro  $X \sim N(\mu, \sigma^2)$  je

$$E(Kurt(X)) = -\frac{6}{n+1} \quad var(Kurt(X)) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$$

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.

# Statistické charakteristiky

Další důležité výběrové charakteristiky:

- Výběrová šikmost (skewness):  $Skew(X) = \frac{m_3(X)}{m_2^{3/2}(X)}$   
pro  $X \sim N(\mu, \sigma^2)$  je

$$E(Skew(X)) = 0 \quad var(Skew(X)) = \frac{6(n-2)}{(n+1)(n+3)}$$

- Výběrová špičatost (kurtosis):  $Kurt(X) = \frac{m_4(X)}{m_2^2(X)} - 3$   
pro  $X \sim N(\mu, \sigma^2)$  je

$$E(Kurt(X)) = -\frac{6}{n+1} \quad var(Kurt(X)) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$$

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.



# Pořadové statistiky

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .



# Pořadové statistiky

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .

- *Uspořádaný výběr*:  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  vznikne z původního výběru  $X_1, X_2, \dots, X_n$  upořádáním podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ .

# Pořadové statistiky

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .

- *Uspořádaný výběr:*  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  vznikne z původního výběru  $X_1, X_2, \dots, X_n$  upořádáním podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ .
- *Pořadová statistika:*  $X_{(k)}$  je náhodná veličina  $X_m$ , která je  $k$ -tá v pořadí podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ . Index  $k$  nazýváme *pořadím veličiny*  $X_m$  a zapisujeme to  $R_m = k$ .

# Pořadové statistiky

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .

- *Uspořádaný výběr:*  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  vznikne z původního výběru  $X_1, X_2, \dots, X_n$  upořádáním podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ .
- *Pořadová statistika:*  $X_{(k)}$  je náhodná veličina  $X_m$ , která je  $k$ -tá v pořadí podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ . Index  $k$  nazýváme *pořadím veličiny*  $X_m$  a zapisujeme to  $R_m = k$ .
- Statistika  $X_{(1)}$  se nazývá minimum,  $X_{(n)}$  je maximum

# Pořadové statistiky

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .

- *Uspořádaný výběr*:  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  vznikne z původního výběru  $X_1, X_2, \dots, X_n$  upořádáním podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ .
- *Pořadová statistika*:  $X_{(k)}$  je náhodná veličina  $X_m$ , která je  $k$ -tá v pořadí podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ . Index  $k$  nazýváme *pořadím veličiny*  $X_m$  a zapisujeme to  $R_m = k$ .
- Statistika  $X_{(1)}$  se nazývá minimum,  $X_{(n)}$  je maximum
- medián  $\tilde{x}_{50}$ : je-li  $n$  liché, je roven  $X_{([\frac{n}{2}] + 1)}$   
pro  $n$  sudé je roven  $(X_{(n/2)} + X_{(n/2+1)})/2$

# Pořadové statistiky

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .

- *Uspořádaný výběr*:  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  vznikne z původního výběru  $X_1, X_2, \dots, X_n$  upořádáním podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ .
- *Pořadová statistika*:  $X_{(k)}$  je náhodná veličina  $X_m$ , která je  $k$ -tá v pořadí podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ . Index  $k$  nazýváme *pořadím veličiny*  $X_m$  a zapisujeme to  $R_m = k$ .
- Statistika  $X_{(1)}$  se nazývá minimum,  $X_{(n)}$  je maximum
- medián  $\tilde{x}_{50}$ : je-li  $n$  liché, je roven  $X_{([\frac{n}{2}] + 1)}$   
pro  $n$  sudé je roven  $(X_{(n/2)} + X_{(n/2+1)})/2$
- dolní kvartil  $\tilde{x}_{25}$ :  $X_{([\frac{n}{4}] + 1)}$  resp.  $(X_{(n/4)} + X_{(n/4+1)})/2$

# Pořadové statistiky

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .

- *Uspořádaný výběr*:  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  vznikne z původního výběru  $X_1, X_2, \dots, X_n$  upořádáním podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ .
- *Pořadová statistika*:  $X_{(k)}$  je náhodná veličina  $X_m$ , která je  $k$ -tá v pořadí podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ . Index  $k$  nazýváme *pořadím veličiny*  $X_m$  a zapisujeme to  $R_m = k$ .
- Statistika  $X_{(1)}$  se nazývá minimum,  $X_{(n)}$  je maximum
- medián  $\tilde{x}_{50}$ : je-li  $n$  liché, je roven  $X_{([\frac{n}{2}] + 1)}$   
pro  $n$  sudé je roven  $(X_{(n/2)} + X_{(n/2+1)})/2$
- dolní kvartil  $\tilde{x}_{25}$ :  $X_{([\frac{n}{4}] + 1)}$  resp.  $(X_{(n/4)} + X_{(n/4+1)})/2$
- horní kvartil  $\tilde{x}_{75}$ :  $X_{([\frac{3n}{4}] + 1)}$  resp.  $(X_{(3n/4-1)} + X_{(3n/4)})/2$

# Pořadové statistiky

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .

- *Uspořádaný výběr*:  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  vznikne z původního výběru  $X_1, X_2, \dots, X_n$  upořádáním podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ .
- *Pořadová statistika*:  $X_{(k)}$  je náhodná veličina  $X_m$ , která je  $k$ -tá v pořadí podle velikosti pozorovaných hodnot  $x_1, x_2, \dots, x_n$ . Index  $k$  nazýváme *pořadím veličiny*  $X_m$  a zapisujeme to  $R_m = k$ .
- Statistika  $X_{(1)}$  se nazývá minimum,  $X_{(n)}$  je maximum
- medián  $\tilde{x}_{50}$ : je-li  $n$  liché, je roven  $X_{(\lfloor n/2 \rfloor + 1)}$   
pro  $n$  sudé je roven  $(X_{(n/2)} + X_{(n/2+1)})/2$
- dolní kvartil  $\tilde{x}_{25}$ :  $X_{(\lfloor n/4 \rfloor + 1)}$  resp.  $(X_{(n/4)} + X_{(n/4+1)})/2$
- horní kvartil  $\tilde{x}_{75}$ :  $X_{(\lfloor 3n/4 \rfloor + 1)}$  resp.  $(X_{(3n/4-1)} + X_{(3n/4)})/2$

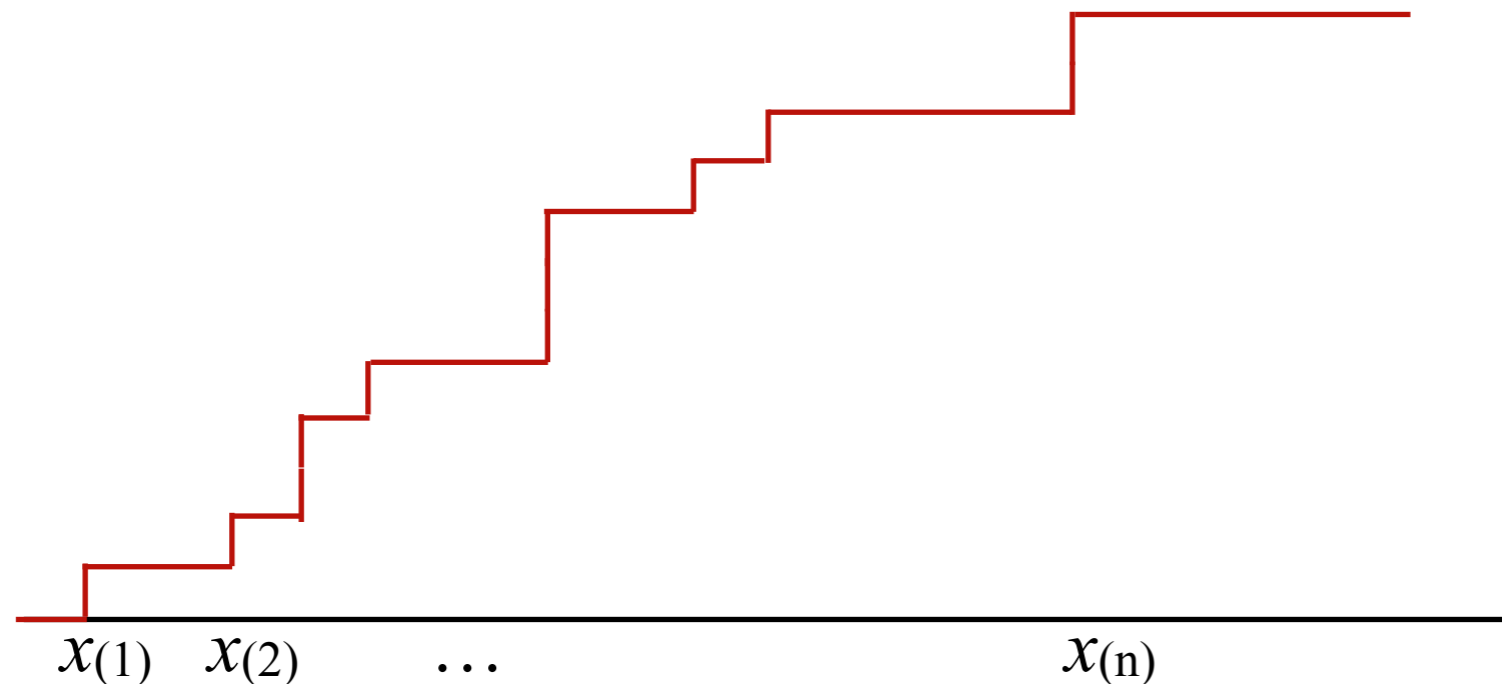




# Empirická distribuční funkce

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .

Empirická distribuční funkce:



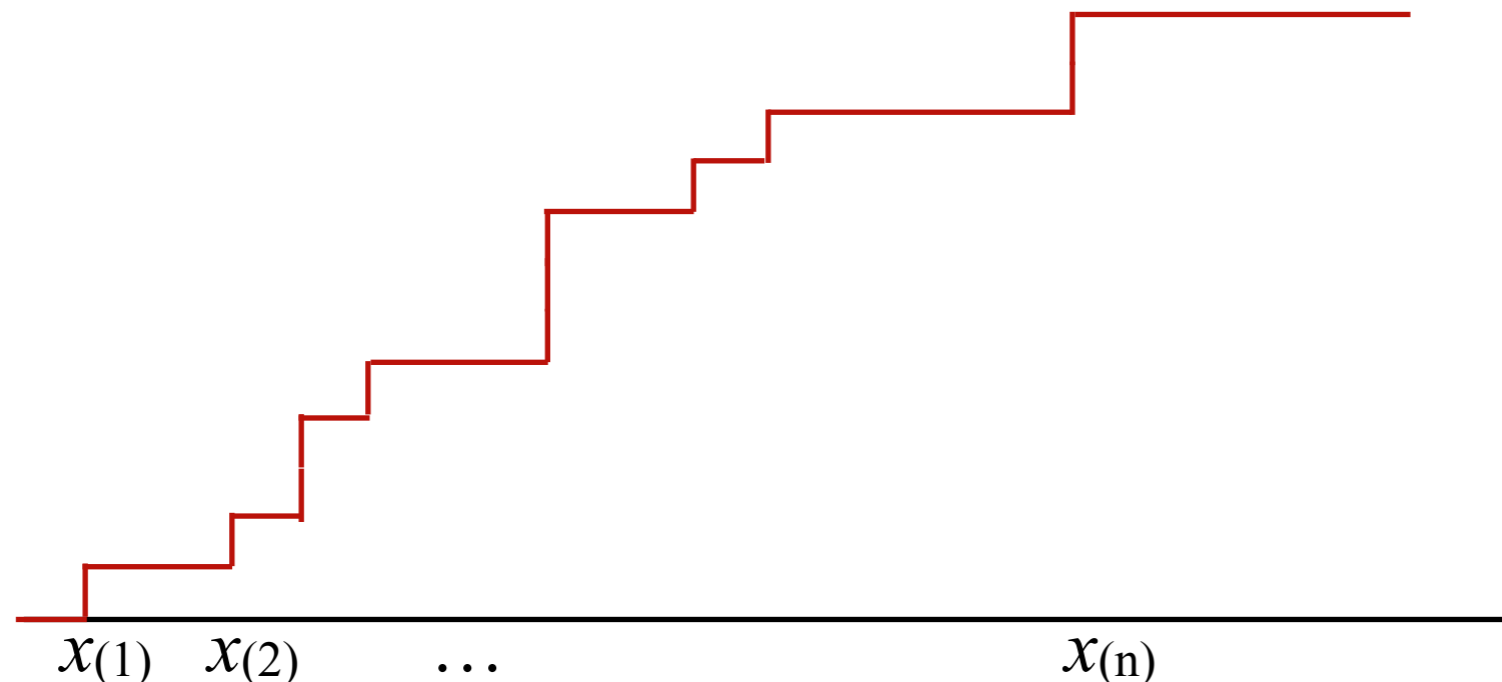
# Empirická distribuční funkce

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .

Empirická distribuční funkce:

vycházíme z uspořádaného výběru:  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ . Potom

$$F_n(x_{(i)}) = \frac{i}{n} \quad \text{a tedy} \quad F_n(x) = \frac{\max\{k : X_{(k)} \leq x\}}{n}, \quad x \in \mathbf{R}$$



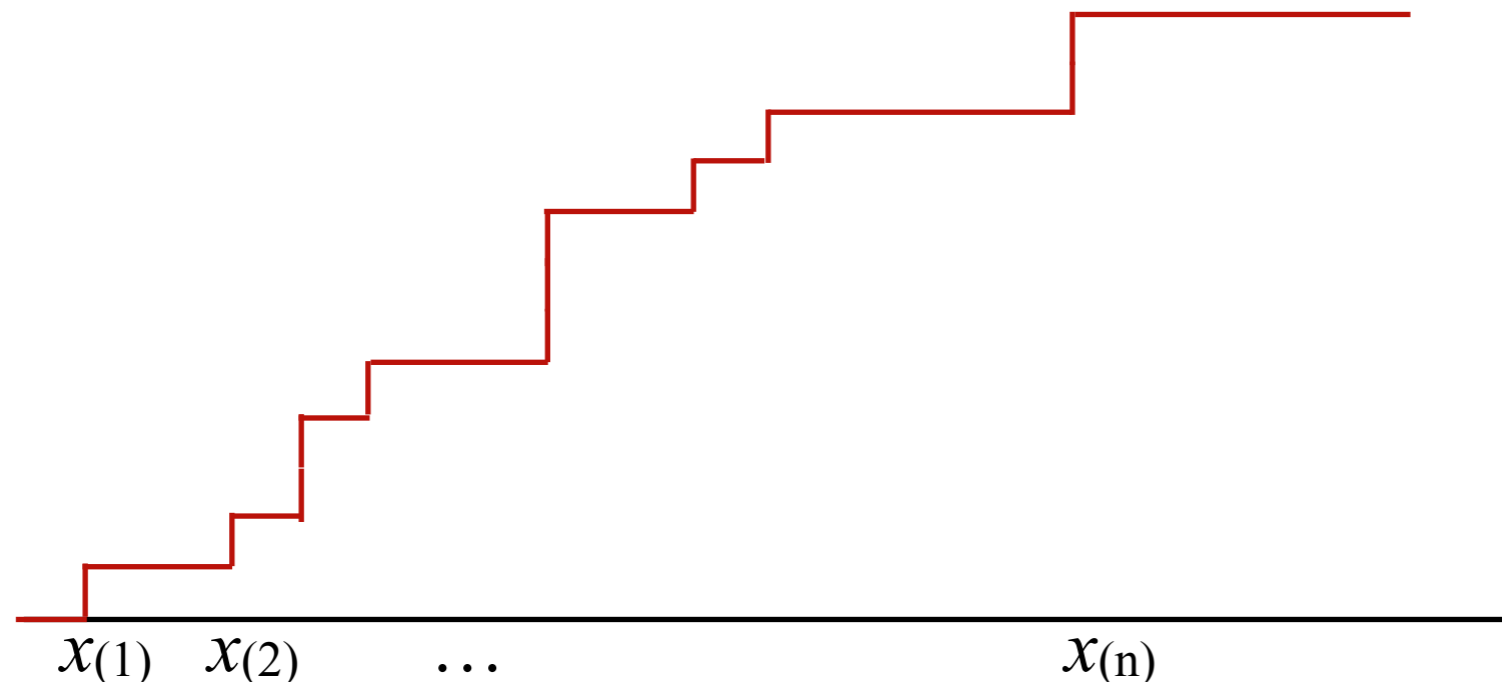
# Empirická distribuční funkce

Statistické charakteristiky: jsou spočteny na základě pozorování  $x_1, x_2, \dots, x_n$  výběru  $X_1, X_2, \dots, X_n$ .

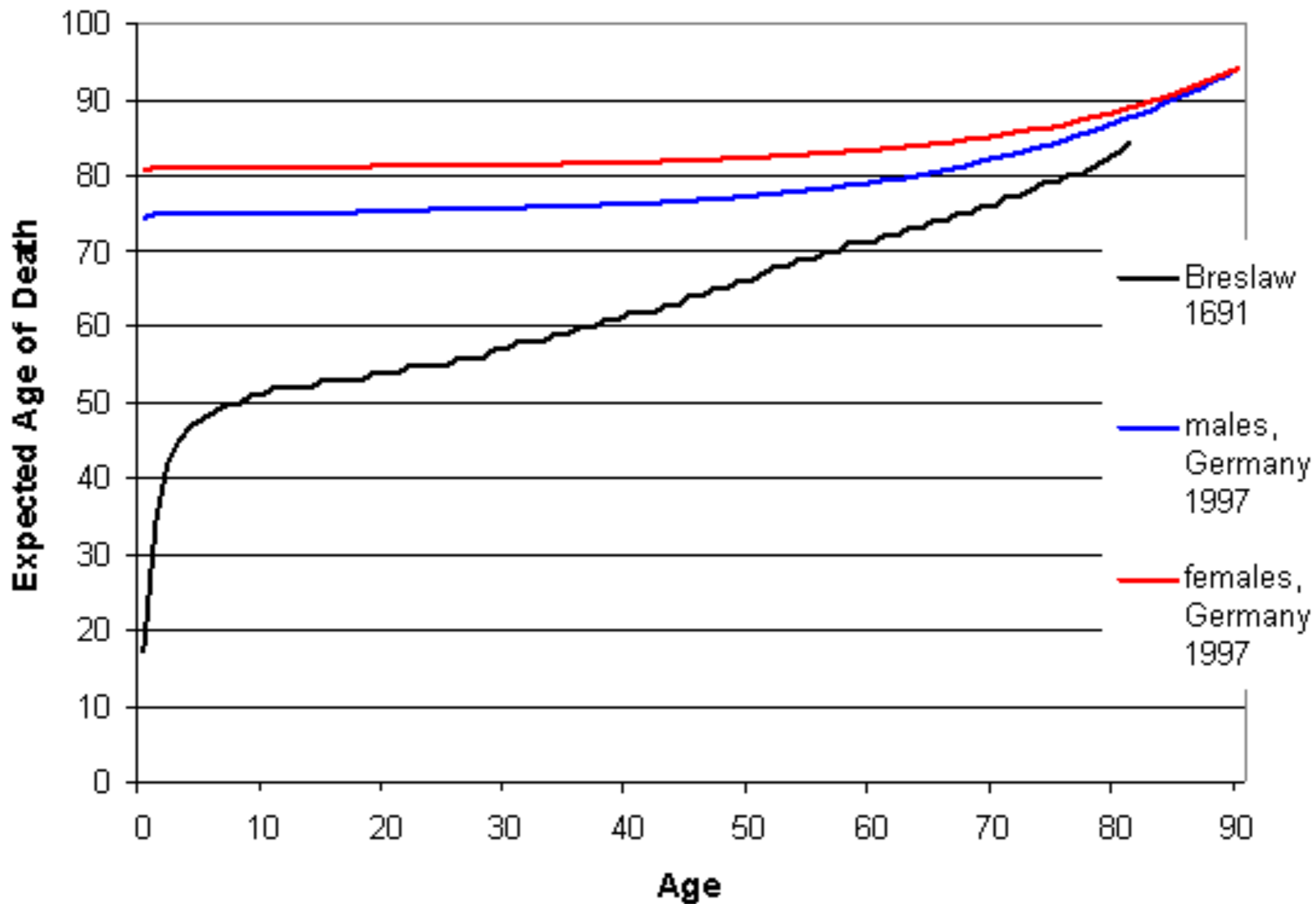
Empirická distribuční funkce:

vycházíme z uspořádaného výběru:  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ . Potom

$$F_n(x_{(i)}) = \frac{i}{n} \quad \text{a tedy} \quad F_n(x) = \frac{\max\{k : X_{(k)} \leq x\}}{n}, \quad x \in \mathbf{R}$$

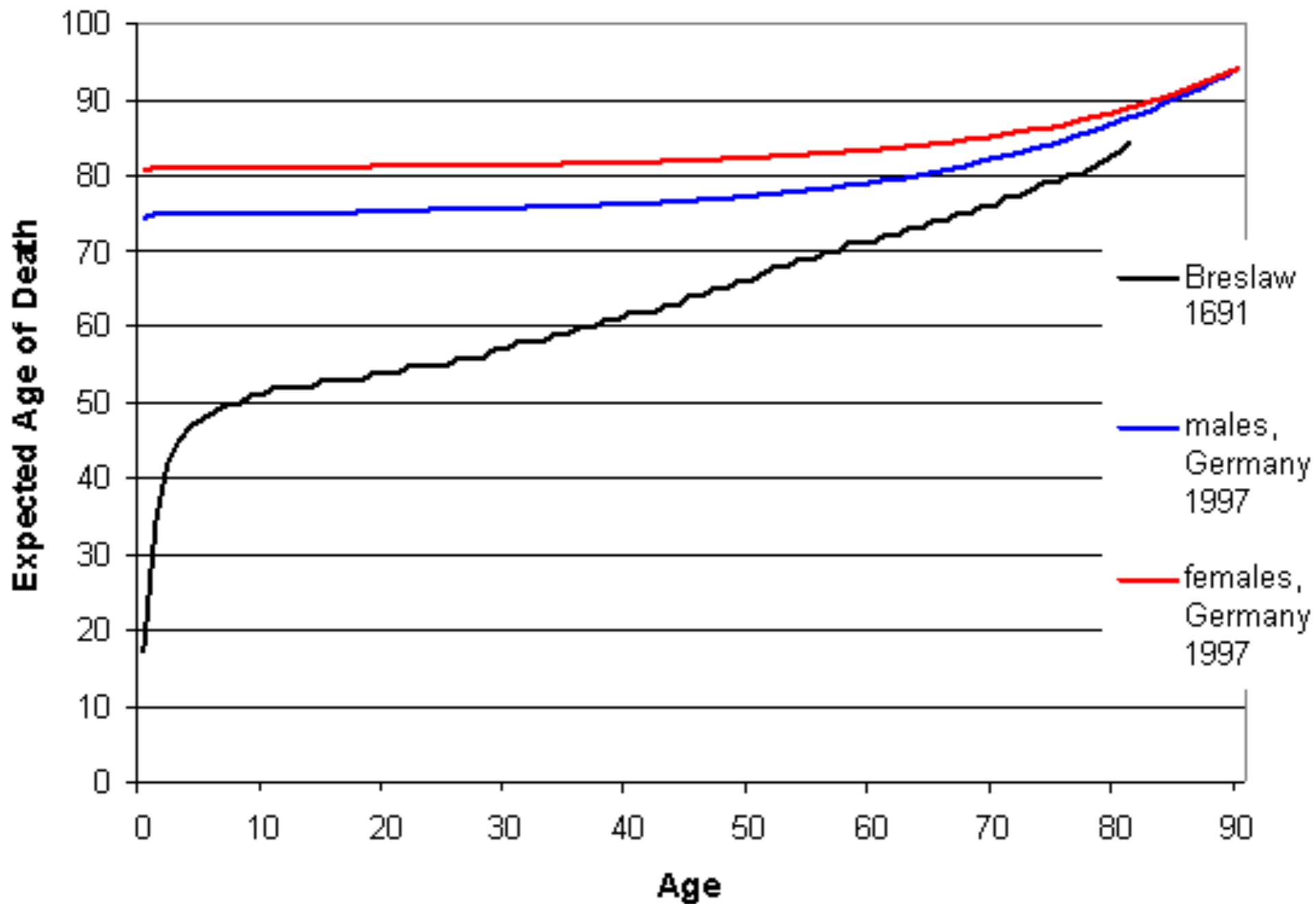


# Empirická distribuční funkce



Data from Edmond Halley's table p.600 and "Durchschnittliche Lebenserwartung im Alter x in Jahren (ex) Deutschland, nach Geschlecht, Sterbetafel 1997/99" made available by [Statistisches Informationssystem GeroStat](#).

# Empirická distribuční funkce

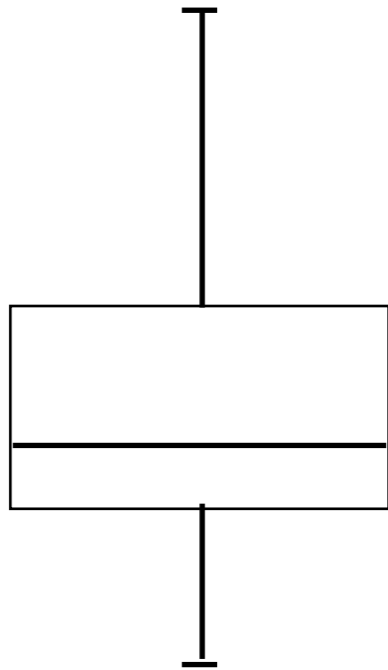


Data from Edmond Halley's table p.600 and "Durchschnittliche Lebenserwartung im Alter x in Jahren (ex) Deutschland, nach Geschlecht, Sterbetafel 1997/99" made available by [Statistisches Informationssystem GeroStat](#).



# Krabicový graf

krabicový graf (Box & Whiskers plot)



# Krabicový graf

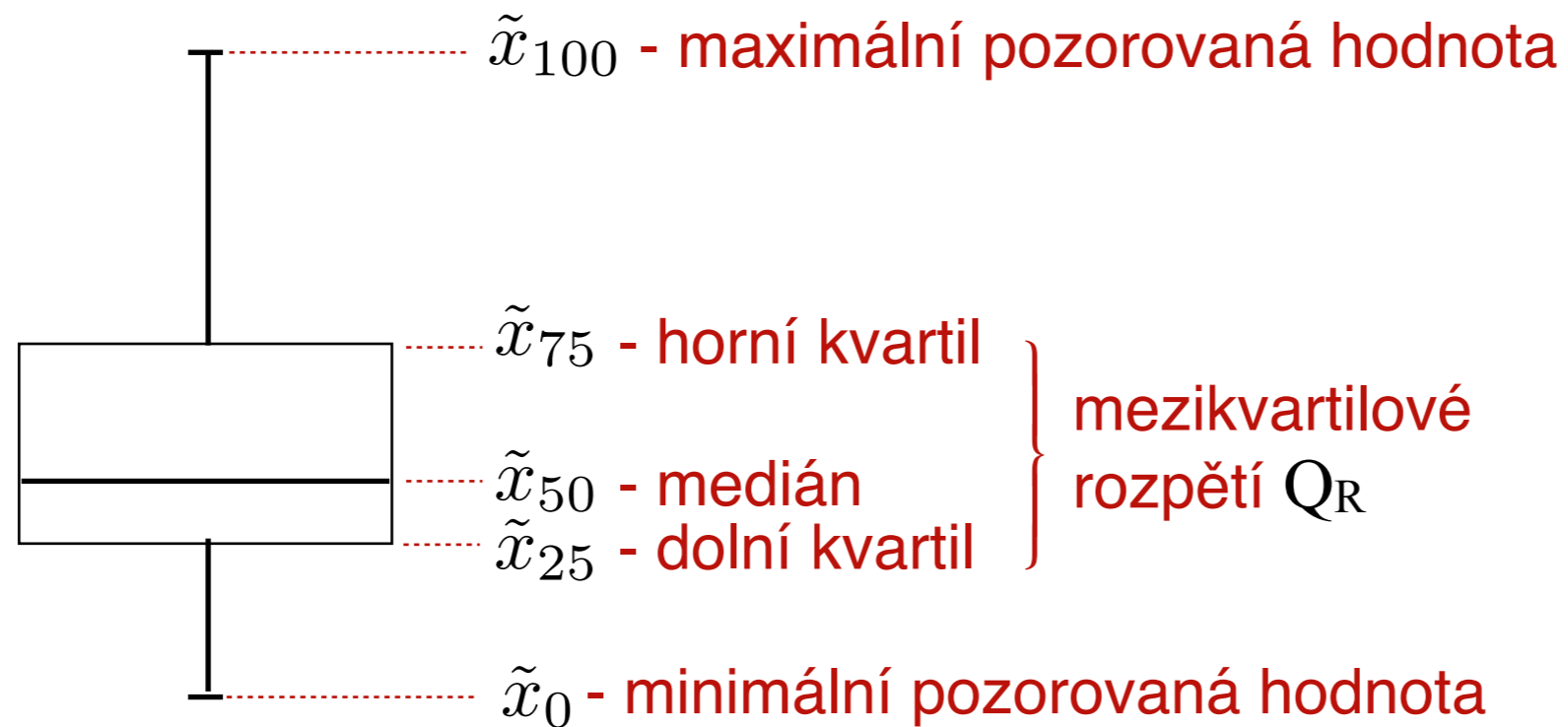
krabicový graf (Box & Whiskers plot)





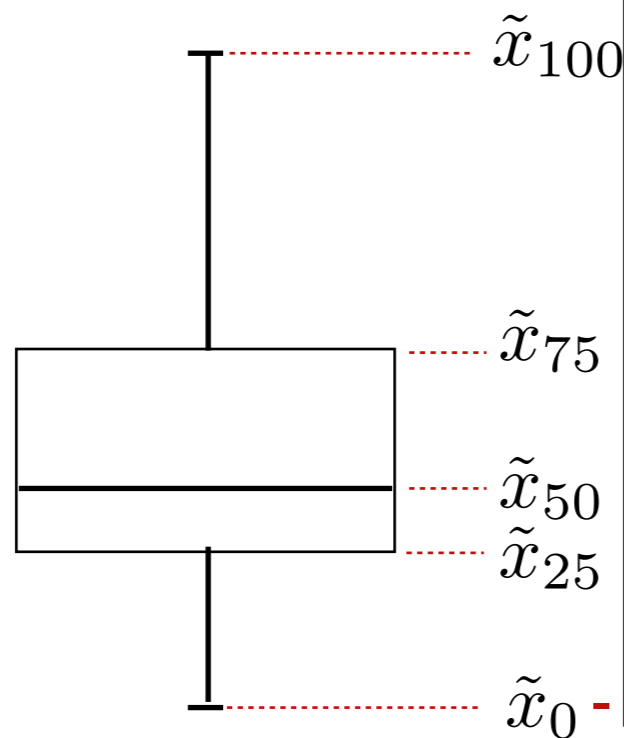
# Krabicový graf

krabicový graf (Box & Whiskers plot)



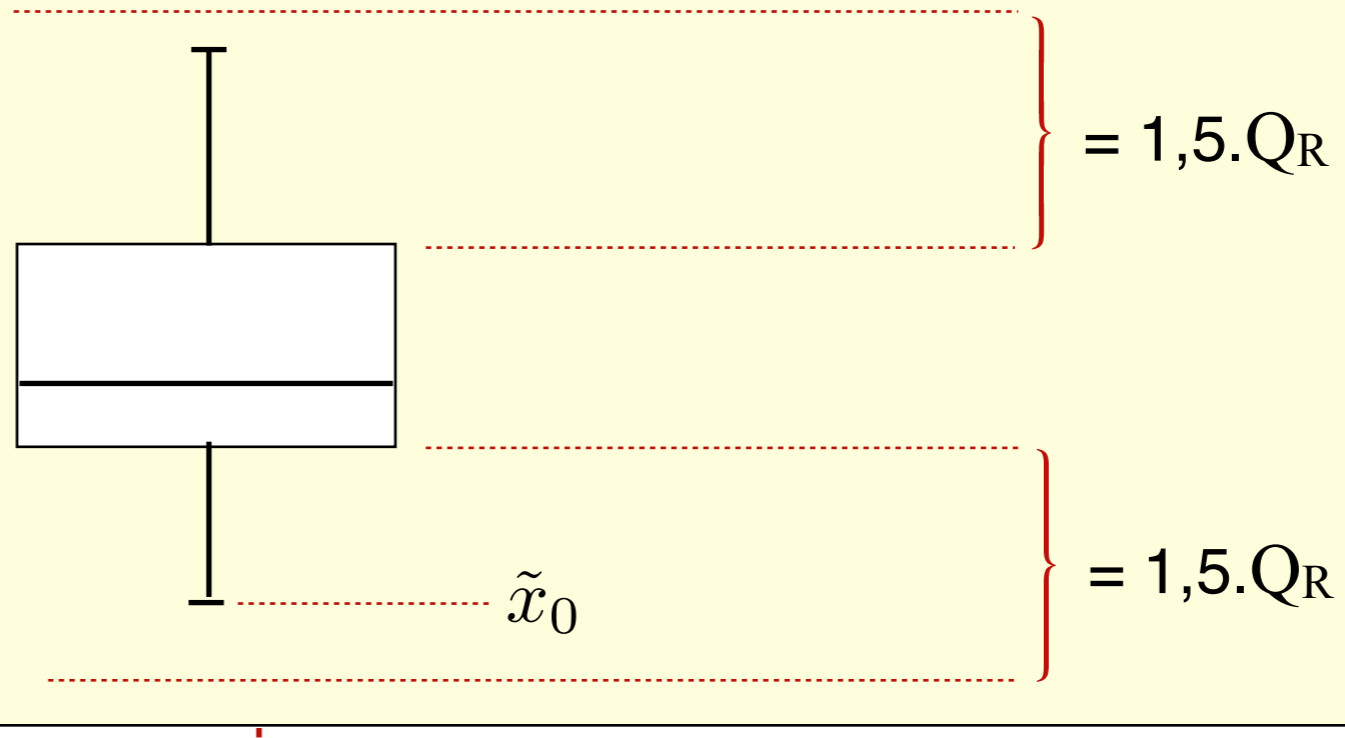
# Krabicový graf

krabicový graf (Box



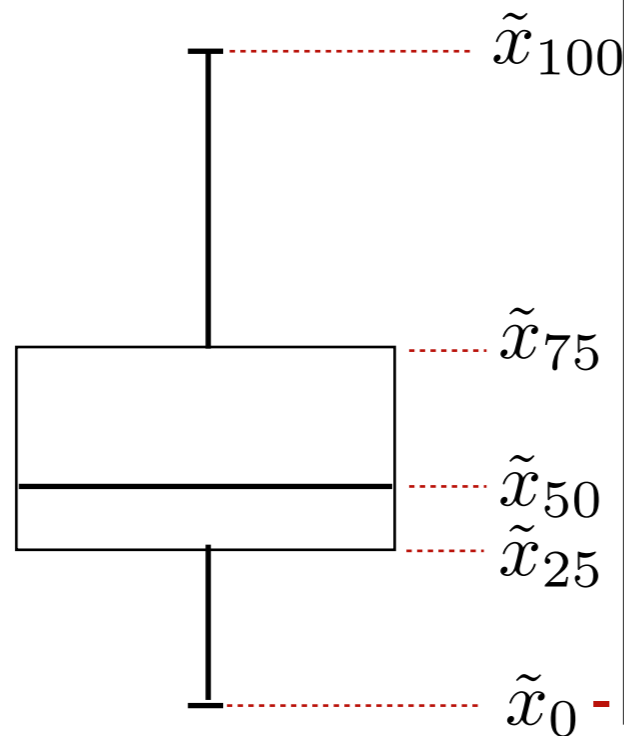
krabicový graf (Box & Whiskers plot)

\*  $\tilde{x}_{100}$  - odlehlé pozorování



# Krabicový graf

krabicový graf (Box



krabicový graf (Box & Whiskers plot)

\*.....  $\tilde{x}_{100}$  - odlehlé pozorování

„vousy“ zde spojují  
napozorovaná data } = 1,5.Q<sub>R</sub>

pouze z intervalu

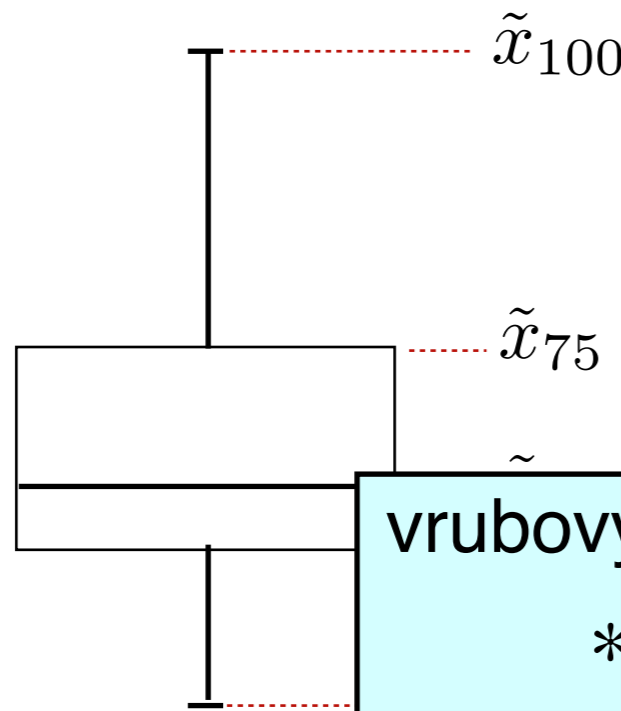
$\langle \tilde{x}_{25} - 1,5Q_R, \tilde{x}_{75} + 1,5Q_R \rangle$

} = 1,5.Q<sub>R</sub>

$\tilde{x}_0$

# Krabicový graf

krabicový graf (Box



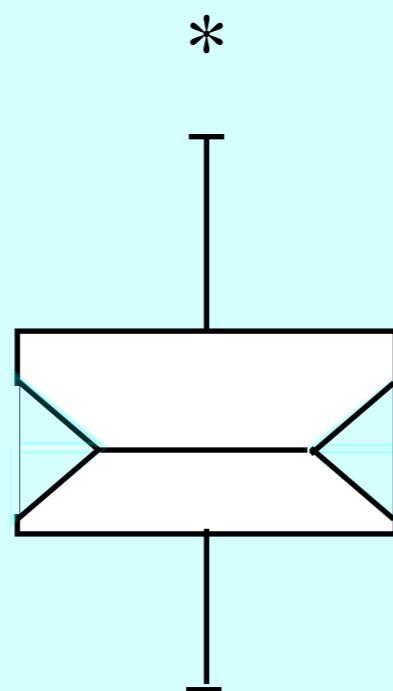
krabicový graf (Box & Whiskers plot)

\* .....  $\tilde{x}_{100}$  - odlehlé pozorování

„vousy“ zde spojují  
napozorovaná data } =  $1,5 \cdot Q_R$

pouze z intervalu  
 $\langle \tilde{x}_{25} - 1,5Q_R, \tilde{x}_{75} + 1,5Q_R \rangle$

vrubový krabicový graf (notched Box & Whiskers plot)

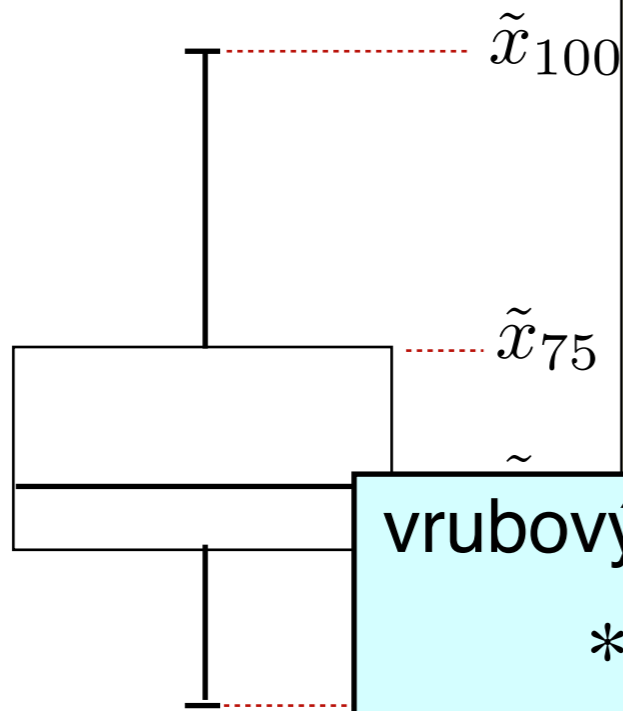


„vrub“ reprezentuje 95% interval  
spolehlivosti pro medián.

$$\tilde{x}_{50} \pm \frac{1,58Q_R}{\sqrt{n}}$$

# Krabicový graf

krabicový graf (Box & Whiskers plot)



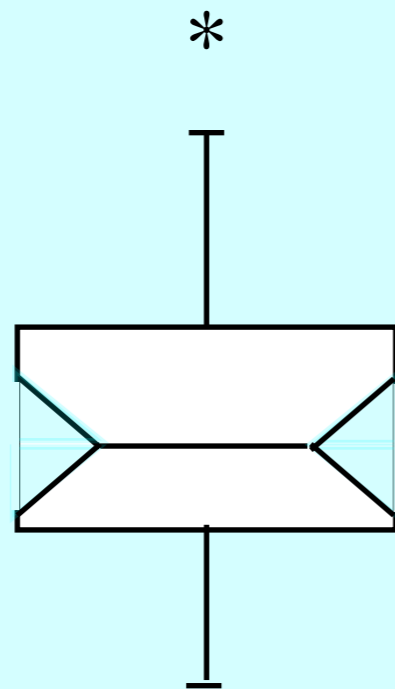
krabicový graf (Box & Whiskers plot)

\* .....  $\tilde{x}_{100}$  - odlehlé pozorování

„vousy“ zde spojují  
napozorovaná data } =  $1,5 \cdot Q_R$

pouze z intervalu  
 $\langle \tilde{x}_{25} - 1,5Q_R, \tilde{x}_{75} + 1,5Q_R \rangle$

vrubový krabicový graf (notched Box & Whiskers plot)



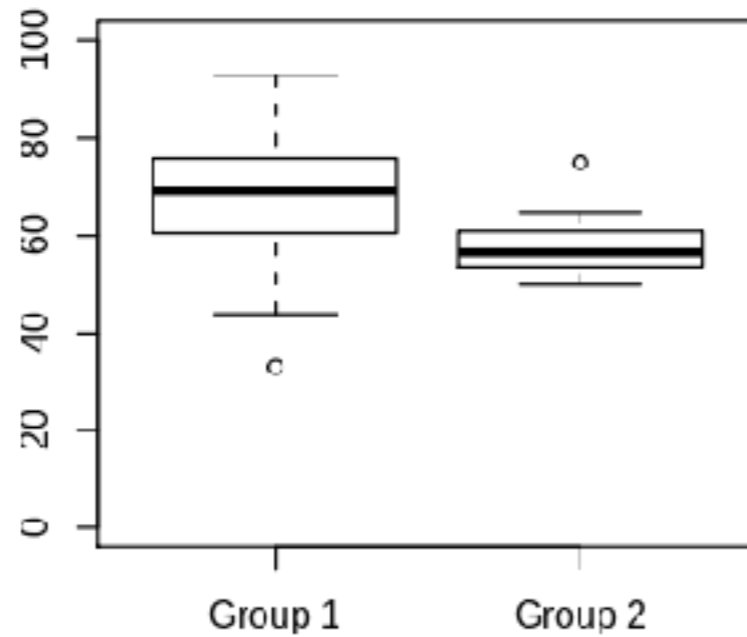
„vrub“ reprezentuje 95% interval  
spolehlivosti pro medián.

$$\tilde{x}_{50} \pm \frac{1,58Q_R}{\sqrt{n}}$$

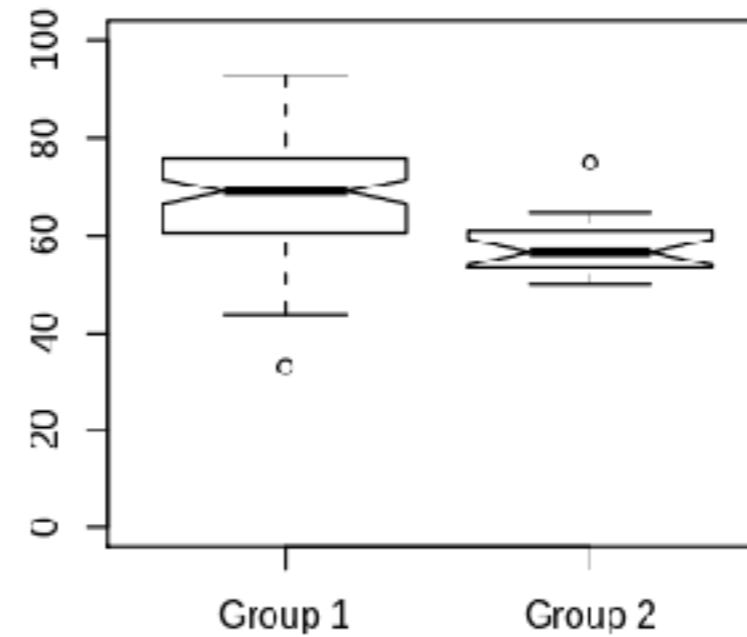


# Krabicový graf

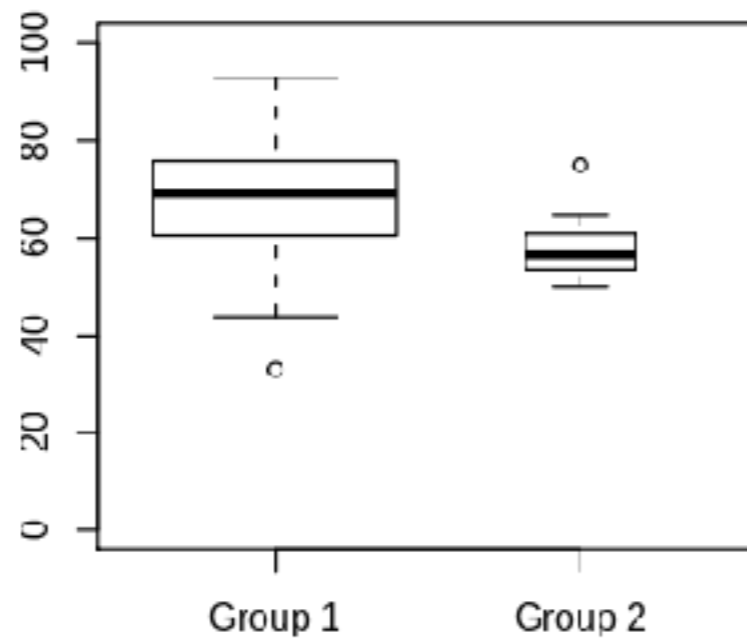
**Traditional Box Plot**



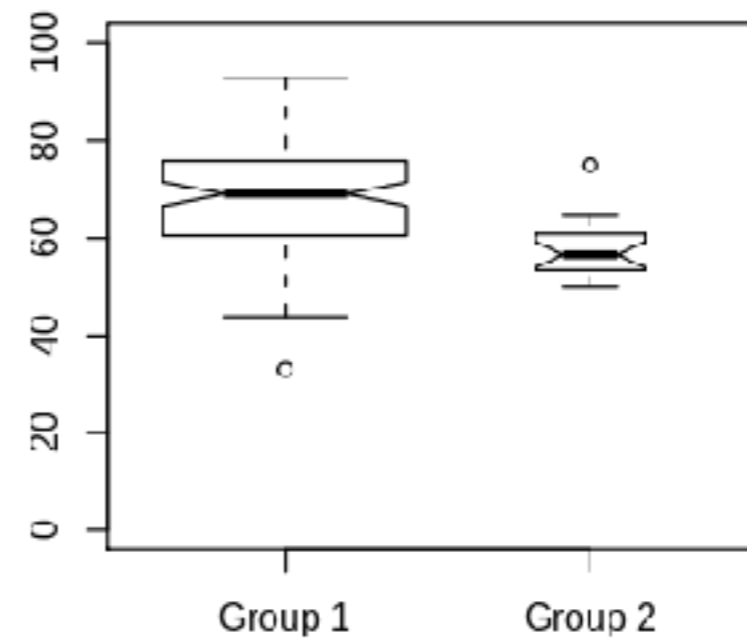
**Notched Box Plot**



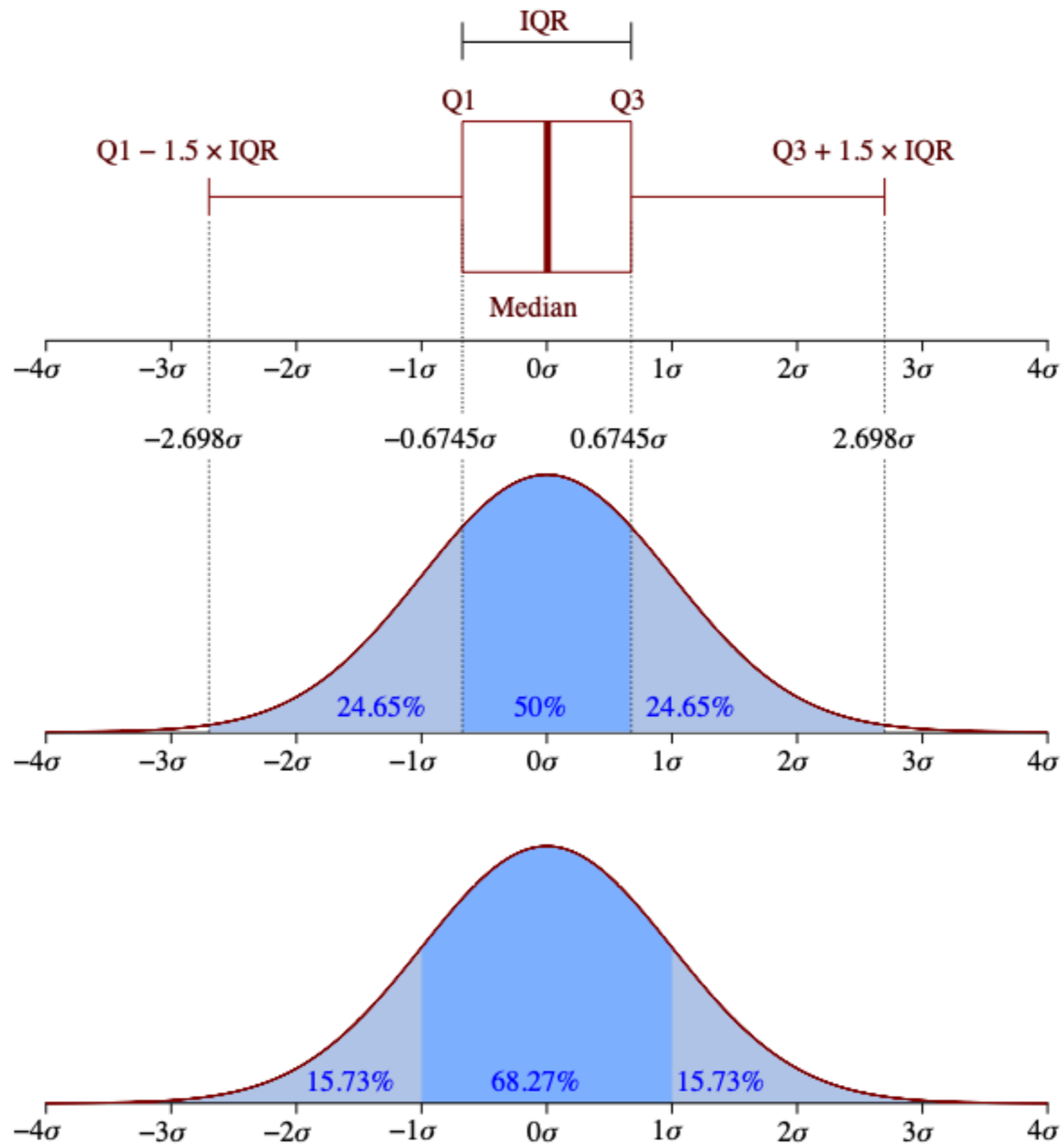
**Variable Width Box Plot**



**Variable Width Notched Box Plot**

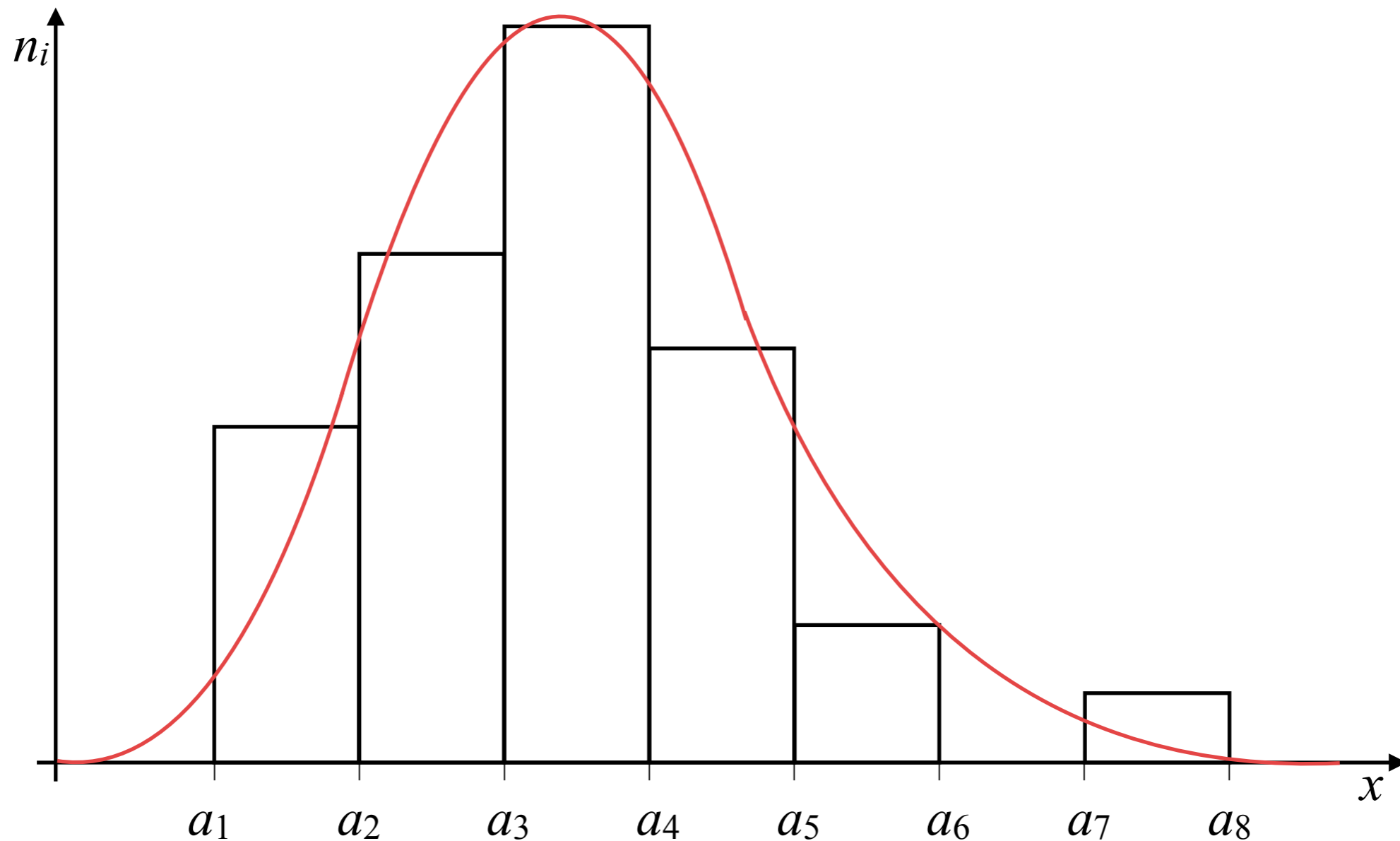


# Krabicový graf



# Frekvenční analýza

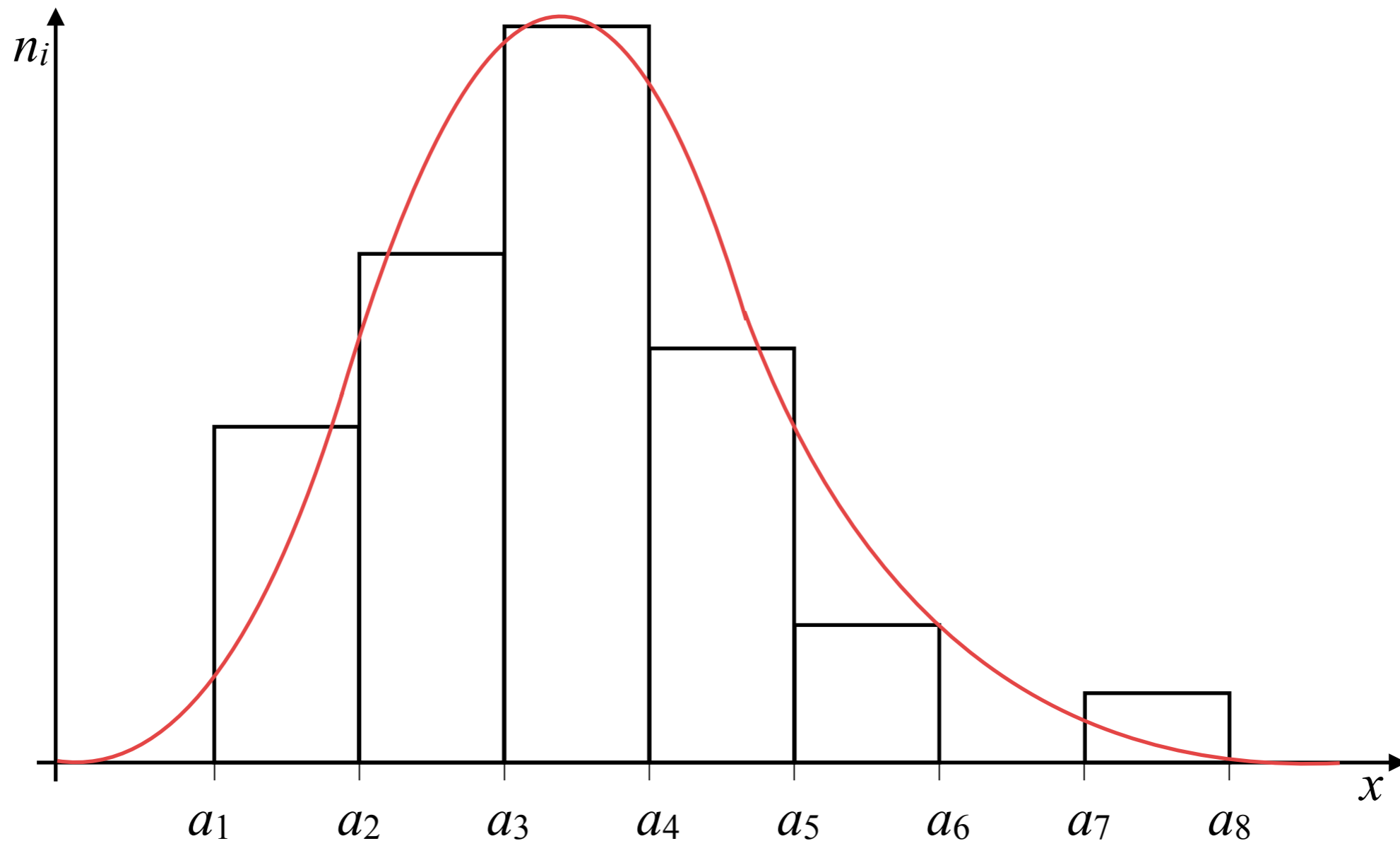
Máme pozorování  $x_1, x_2, \dots, x_n$  náhodného výběru  $X_1, X_2, \dots, X_n$ .





# Frekvenční analýza

Máme pozorování  $x_1, x_2, \dots, x_n$  náhodného výběru  $X_1, X_2, \dots, X_n$ .



# Frekvenční analýza

Máme pozorování  $x_1, x_2, \dots, x_n$  náhodného výběru  $X_1, X_2, \dots, X_n$ .

| pořadí třídy | třídní intervaly | (prosté) absolutní četnosti | (prosté) relativní četnosti | kumulativní četnosti     | kumulativní relativní četnosti |
|--------------|------------------|-----------------------------|-----------------------------|--------------------------|--------------------------------|
| 1            | $a_1 - b_1$      | $n_1$                       | $f_1 = n_1/n$               | $c_1 = n_1$              | $d_1 = c_1/n$                  |
| 2            | $a_2 - b_2$      | $n_2$                       | $f_2 = n_2/n$               | $c_2 = n_1 + n_2$        | $d_2 = (c_1 + c_2)/n$          |
| ⋮            | ⋮                | ⋮                           | ⋮                           | $c_j = \sum_{i=1}^j n_i$ | ⋮                              |
| k            | $a_k - b_k$      | $n_k$                       | $f_k = n_k/n$               | $c_k = n$                | $d_k = c_k/n = 1$              |

# Frekvenční analýza

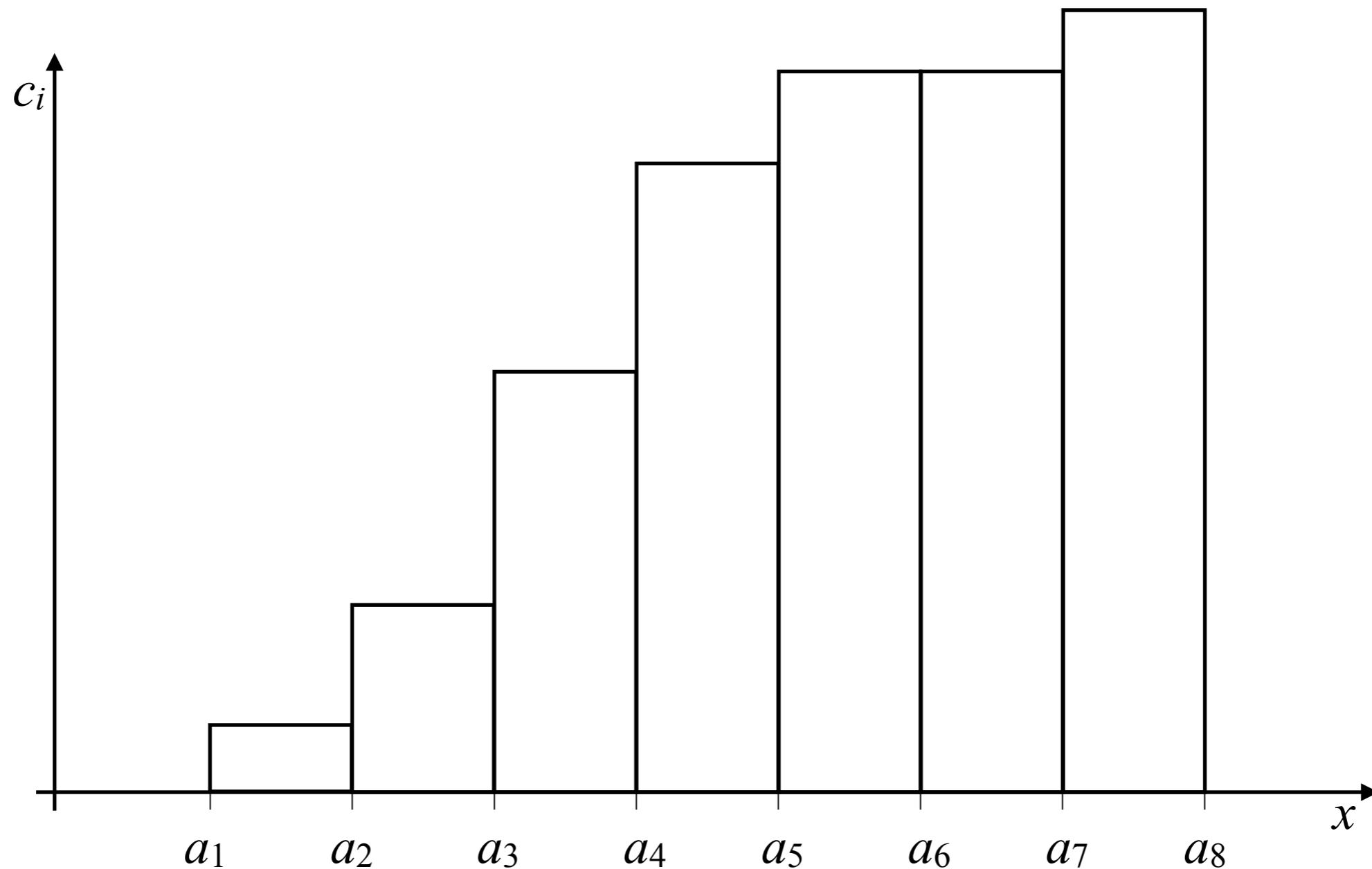
Máme pozorování  $x_1, x_2, \dots, x_n$  náhodného výběru  $X_1, X_2, \dots, X_n$ .

| pořadí třídy | třídní intervaly | (prosté) absolutní četnosti | (prosté) relativní četnosti | kumulativní četnosti     | kumulativní relativní četnosti |
|--------------|------------------|-----------------------------|-----------------------------|--------------------------|--------------------------------|
| 1            | $a_1 - b_1$      | $n_1$                       | $f_1 = n_1/n$               | $c_1 = n_1$              | $d_1 = c_1$                    |
| 2            | $a_2 - b_2$      | $n_2$                       | $f_2 = n_2/n$               | $c_2 = n_1 + n_2$        | $d_2 = (c_1 + c_2)/n$          |
| ⋮            | ⋮                | ⋮                           | ⋮                           | $c_j = \sum_{i=1}^j n_i$ | ⋮                              |
| k            | $a_k - b_k$      | $n_k$                       | $f_k = n_k/n$               | $c_k = n$                | $d_k = c_k/n = 1$              |



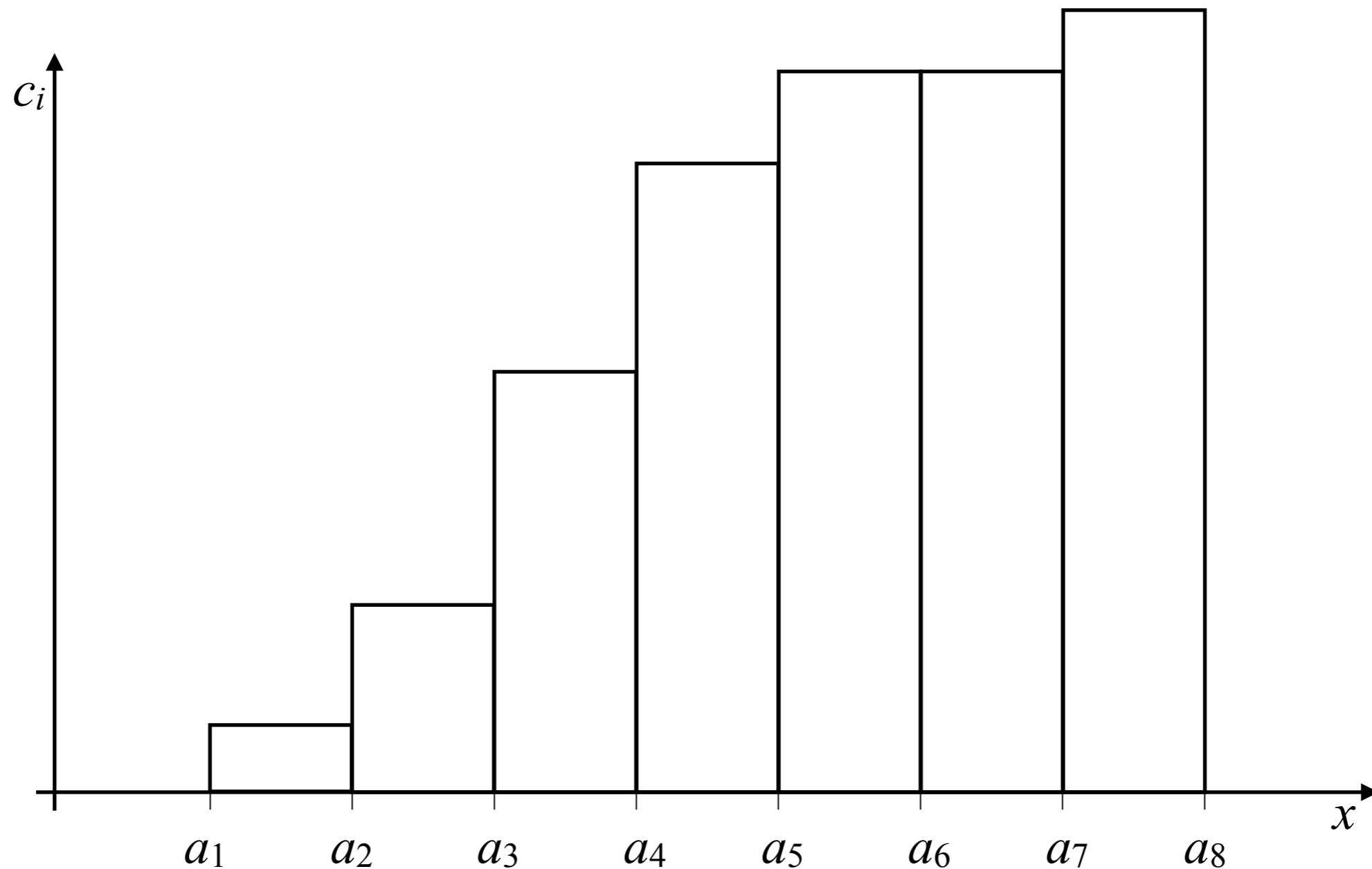
# Frekvenční analýza

Máme pozorování  $x_1, x_2, \dots, x_n$  náhodného výběru  $X_1, X_2, \dots, X_n$ .



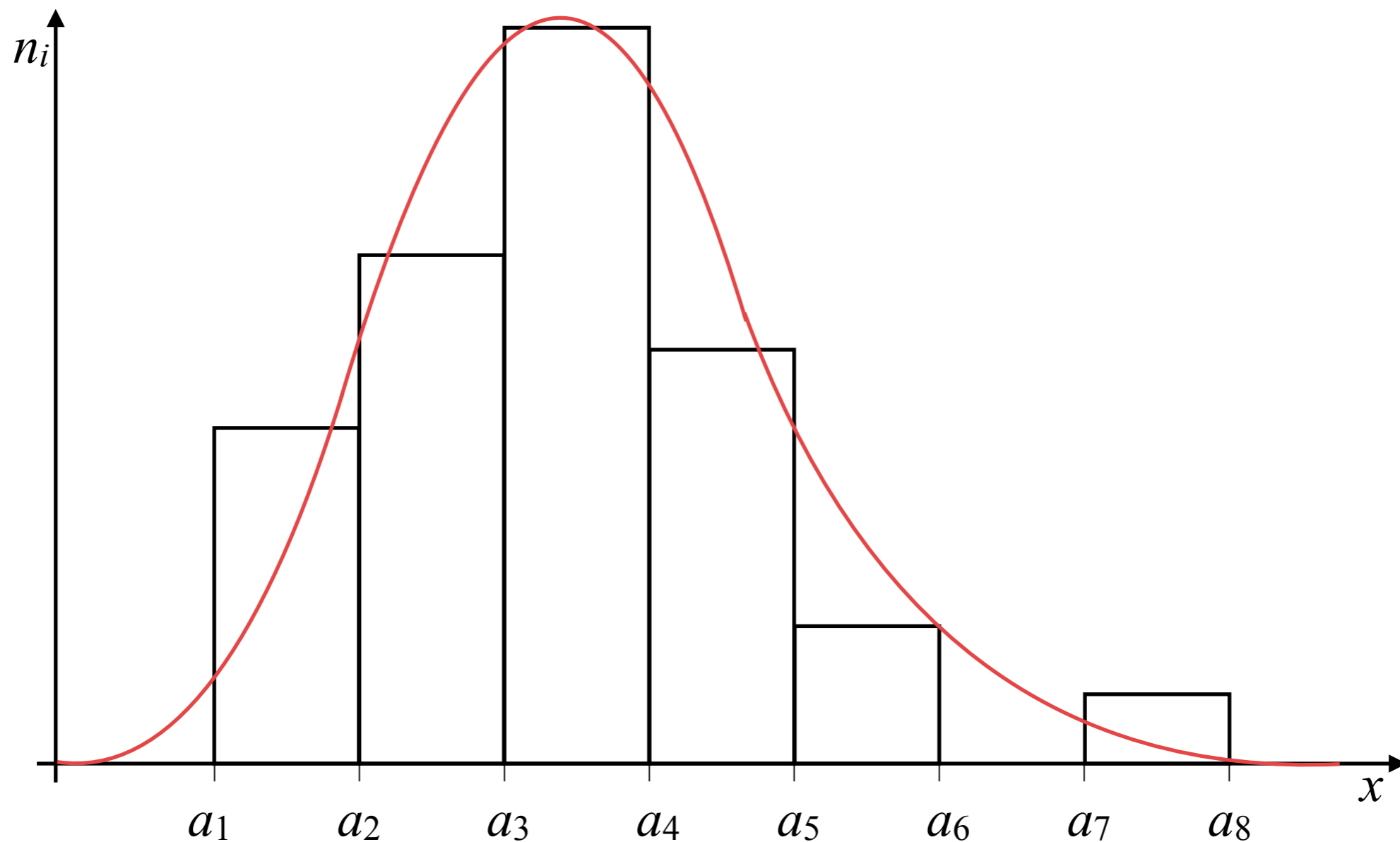
# Frekvenční analýza

Máme pozorování  $x_1, x_2, \dots, x_n$  náhodného výběru  $X_1, X_2, \dots, X_n$ .



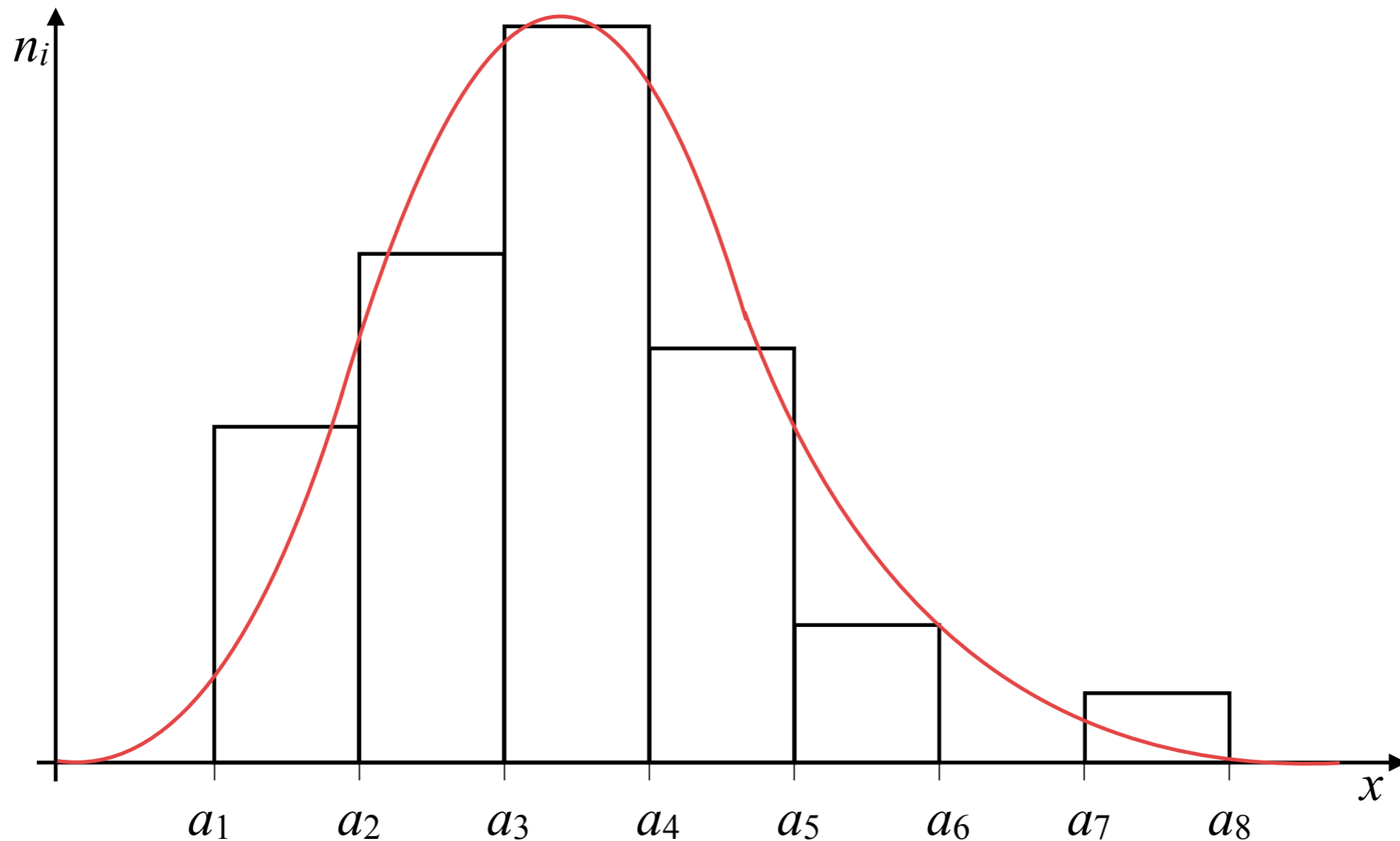
# Frekvenční analýza

Máme pozorování  $x_1, x_2, \dots, x_n$  náhodného výběru  $X_1, X_2, \dots, X_n$ .



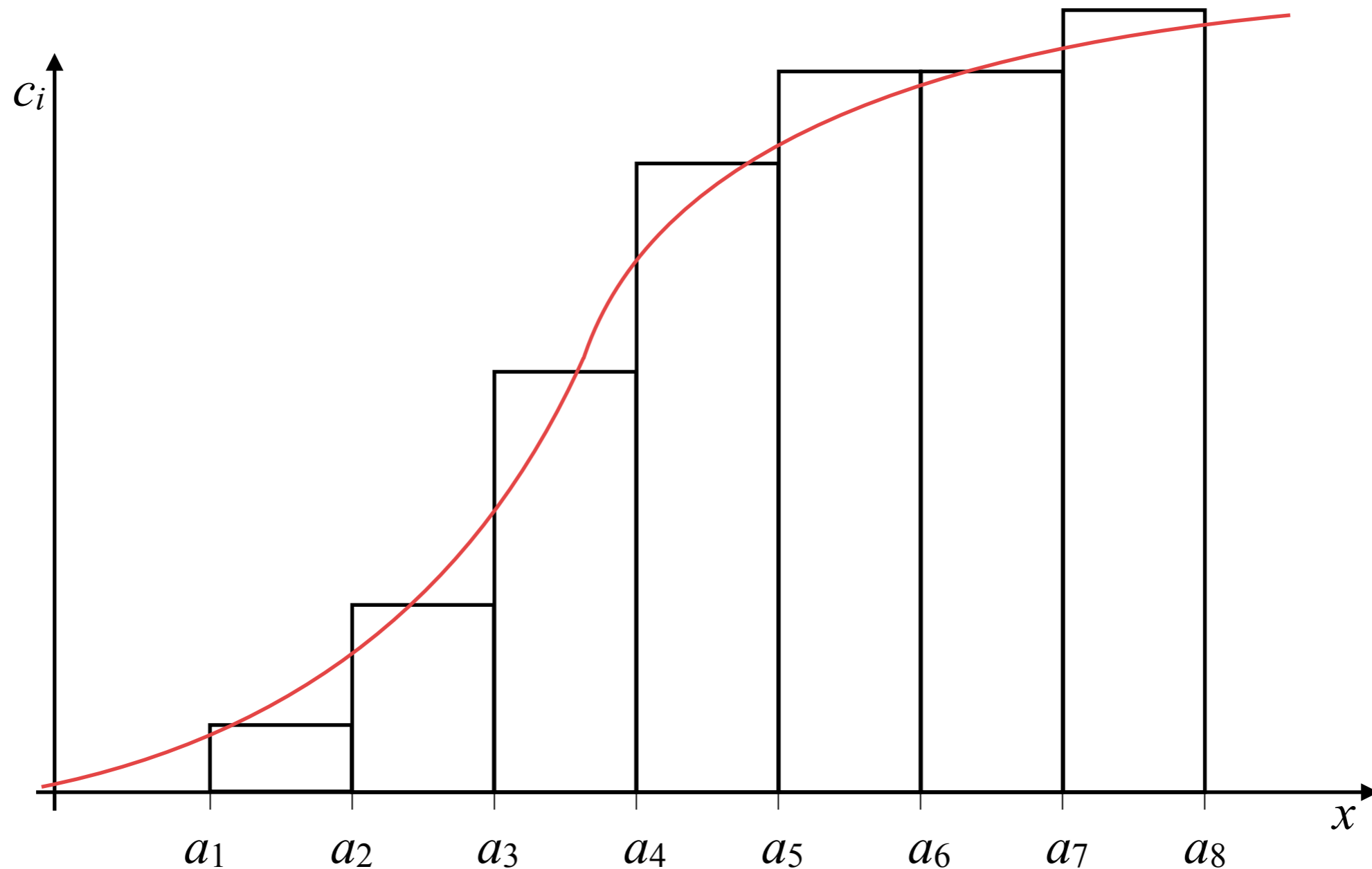
# Frekvenční analýza

Máme pozorování  $x_1, x_2, \dots, x_n$  náhodného výběru  $X_1, X_2, \dots, X_n$ .



# Frekvenční analýza

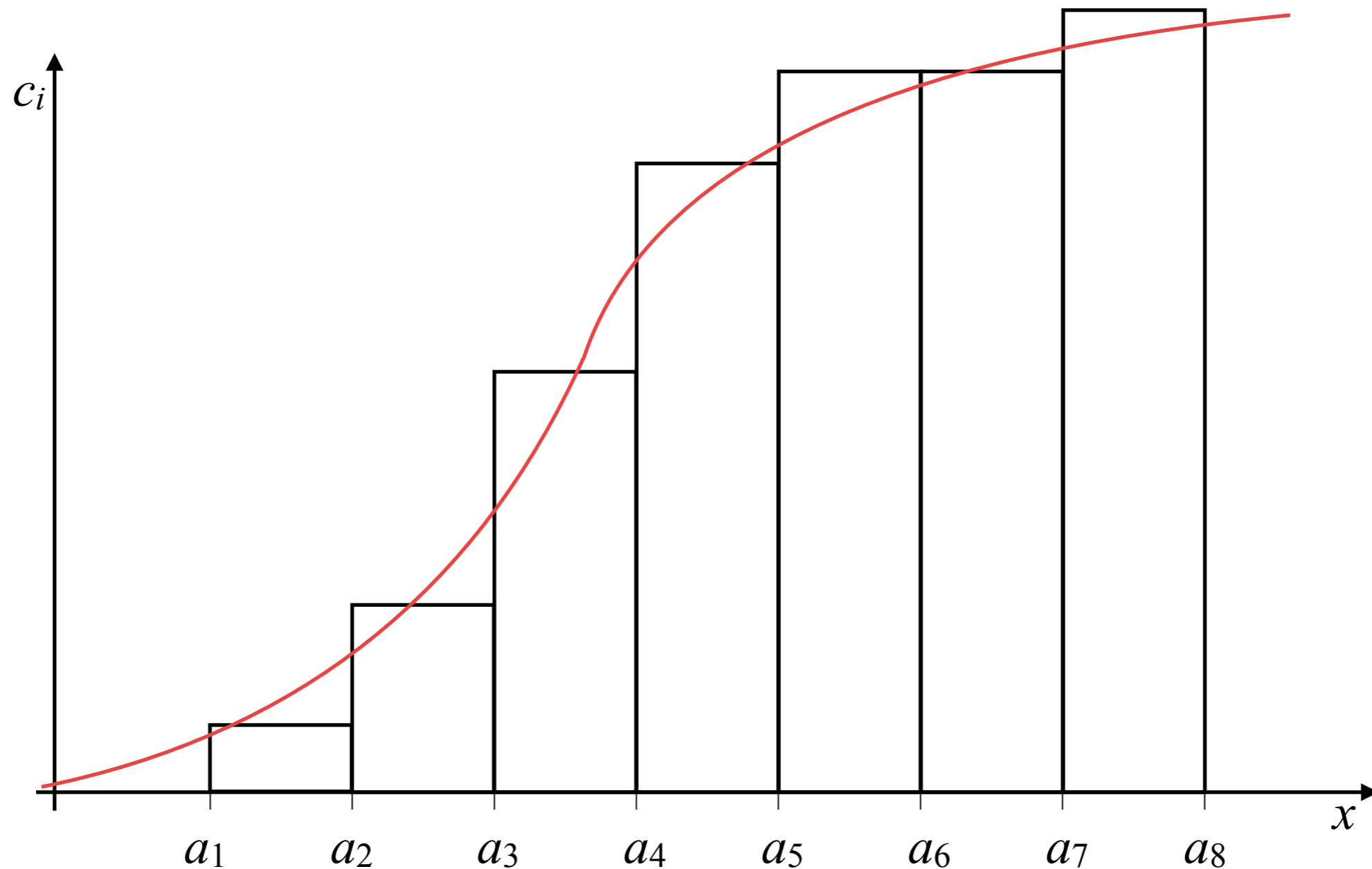
Máme pozorování  $x_1, x_2, \dots, x_n$  náhodného výběru  $X_1, X_2, \dots, X_n$ .



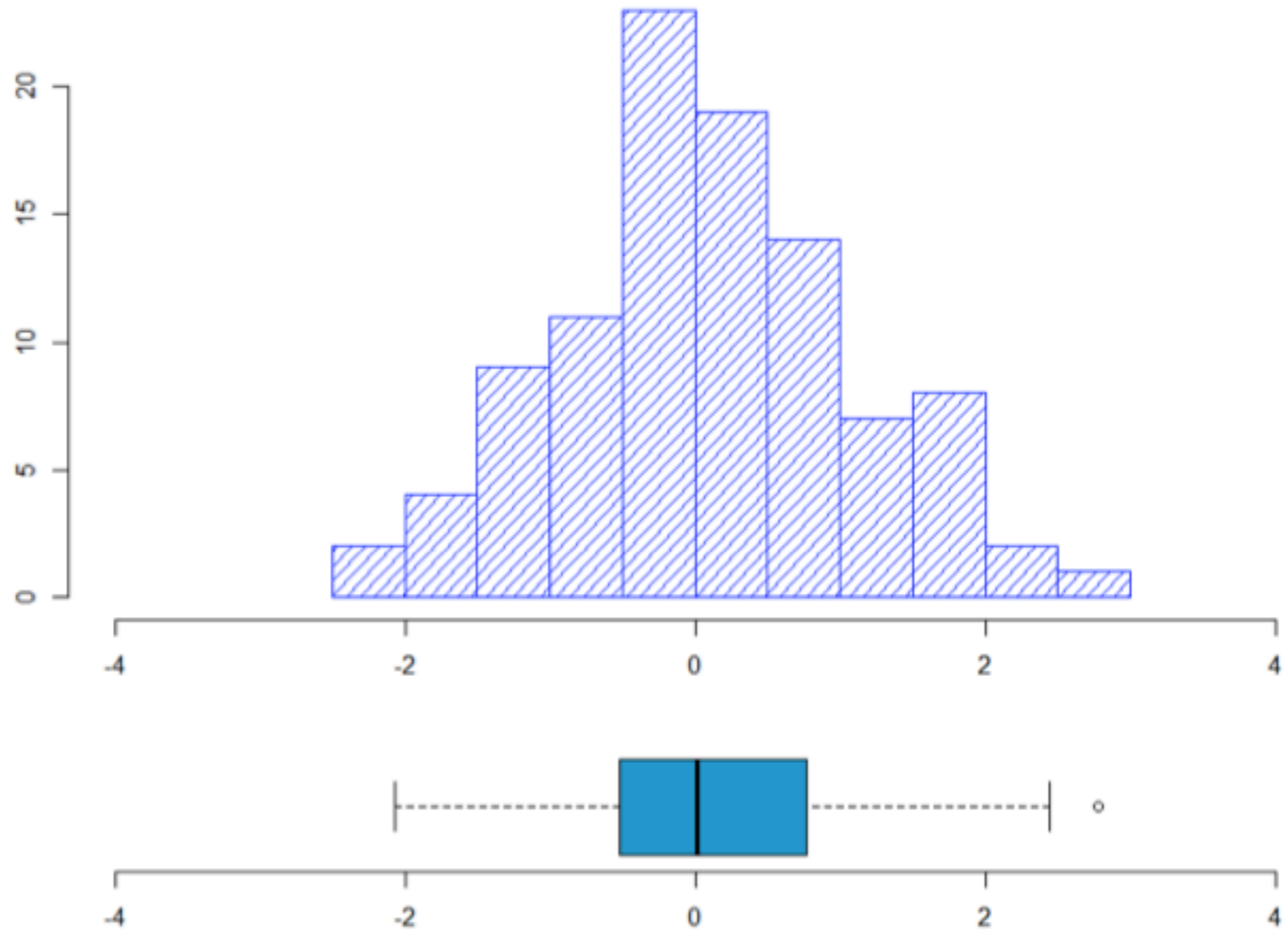


# Frekvenční analýza

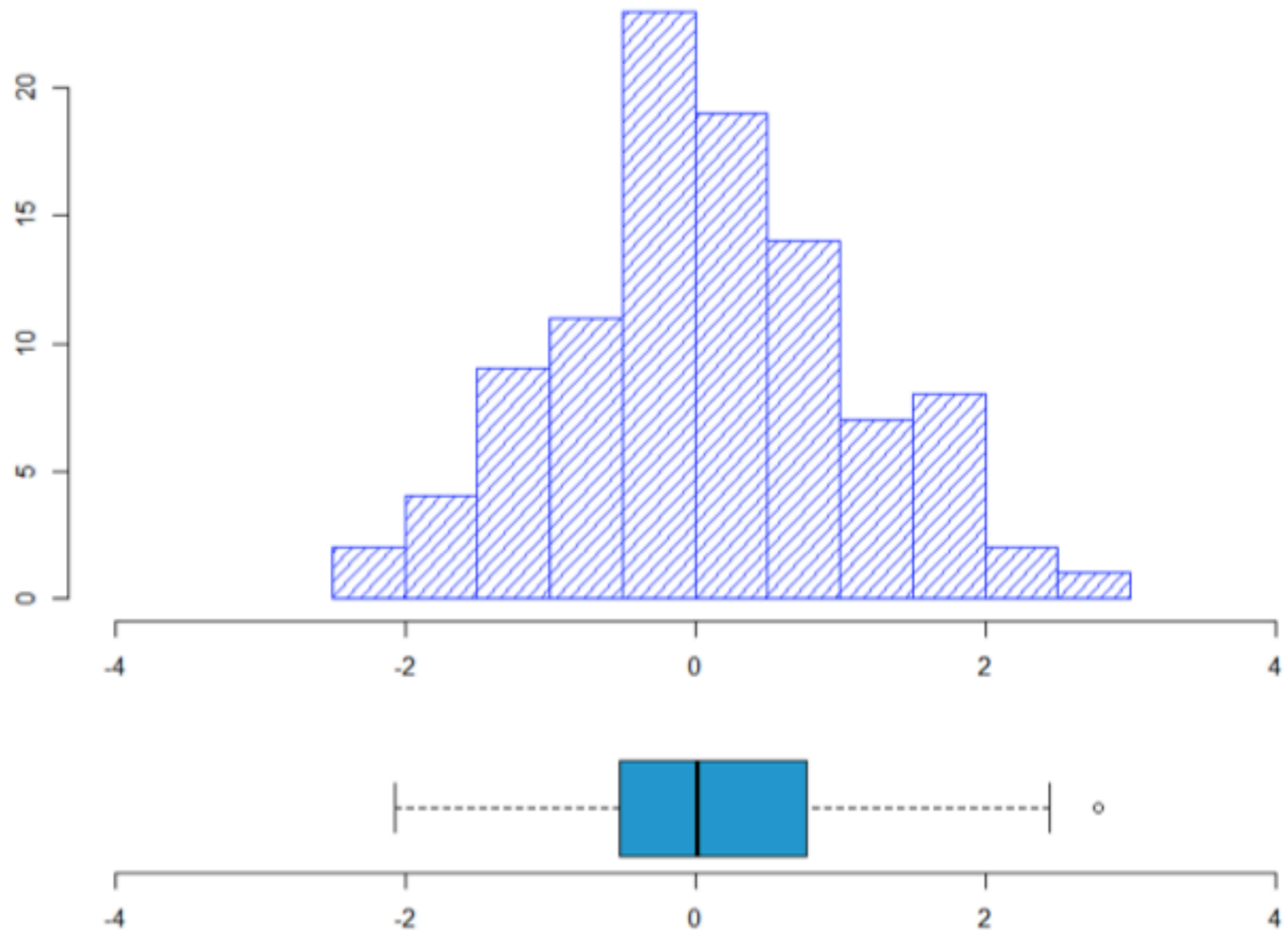
Máme pozorování  $x_1, x_2, \dots, x_n$  náhodného výběru  $X_1, X_2, \dots, X_n$ .



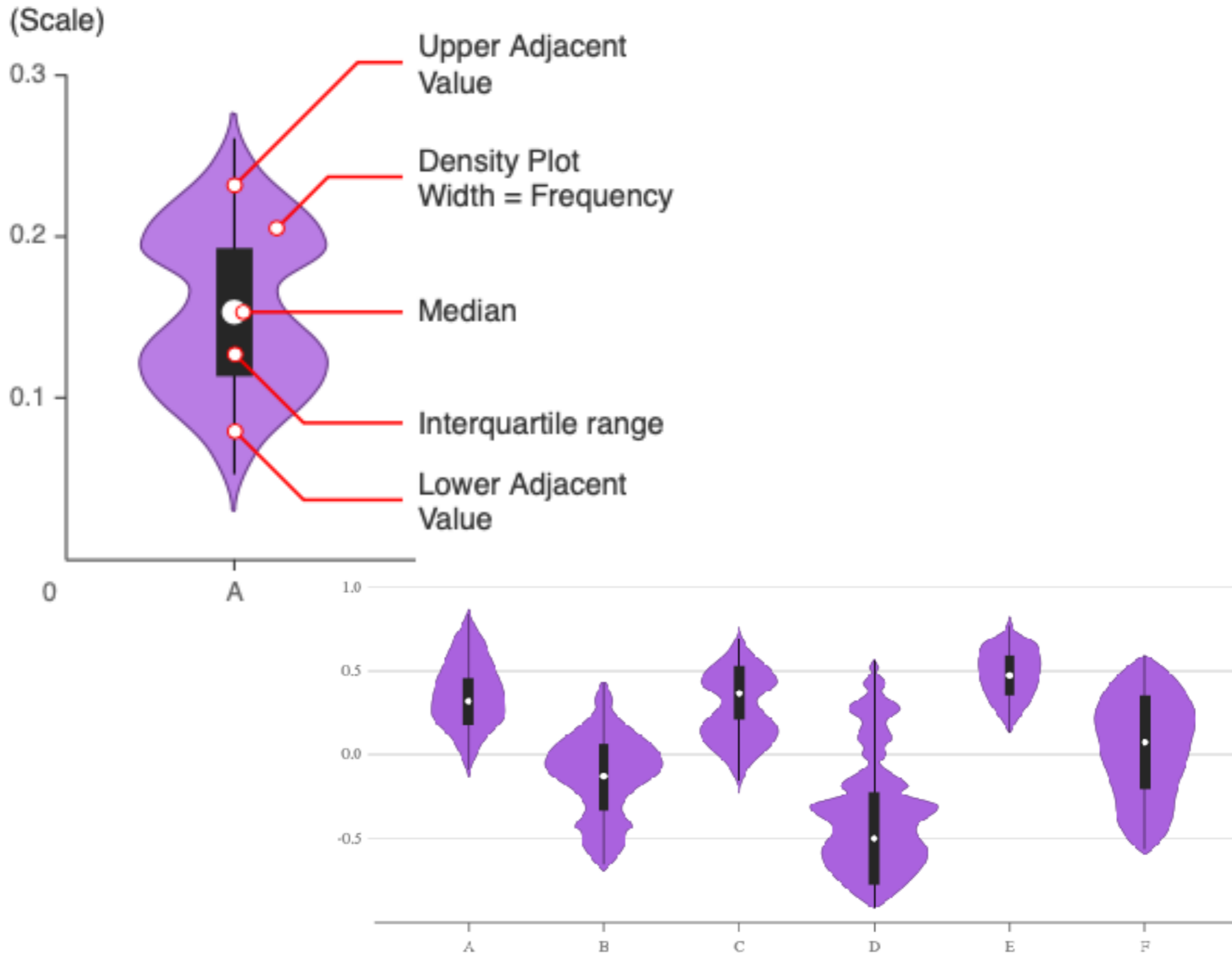
# Histogram x boxplot



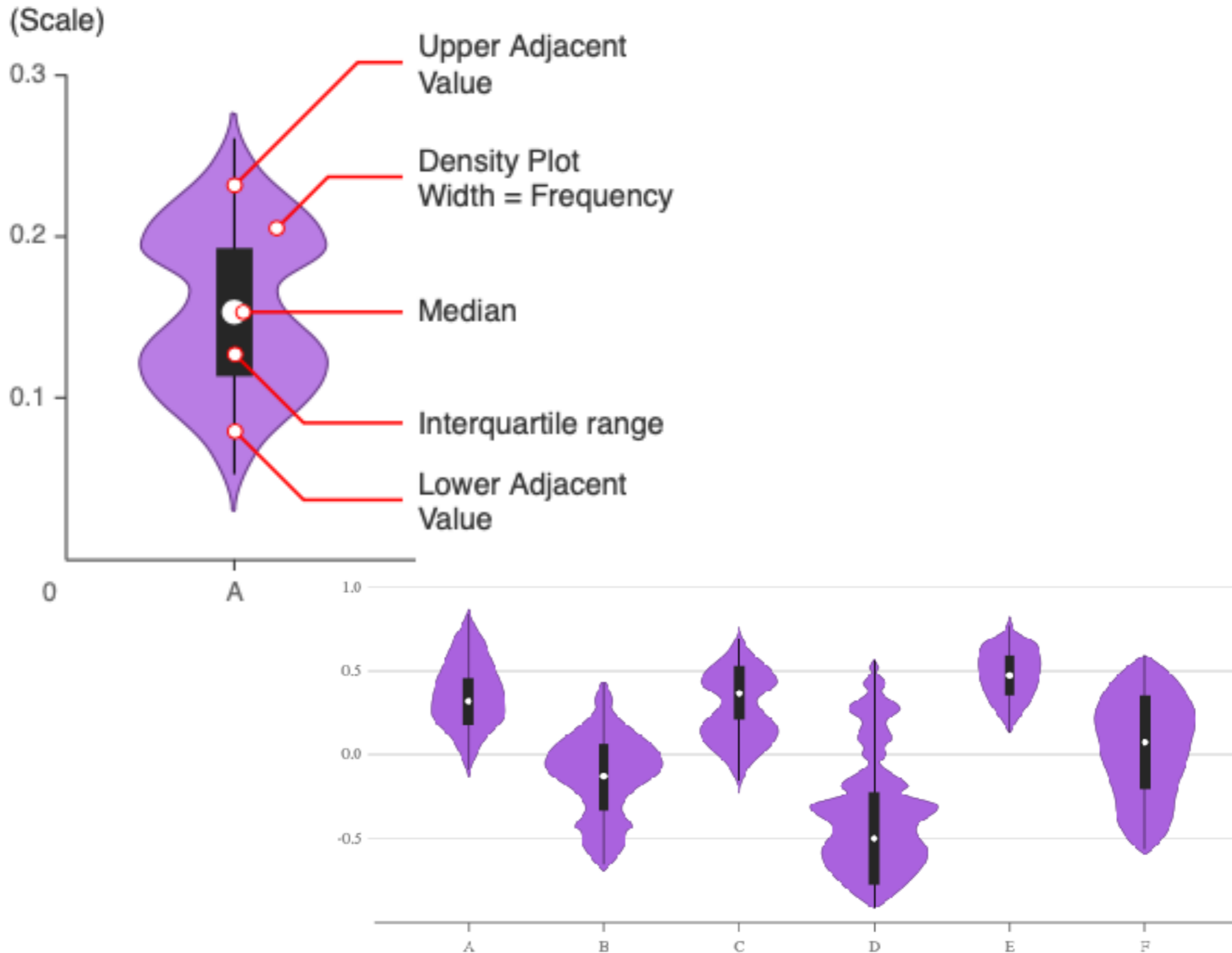
# Histogram x boxplot



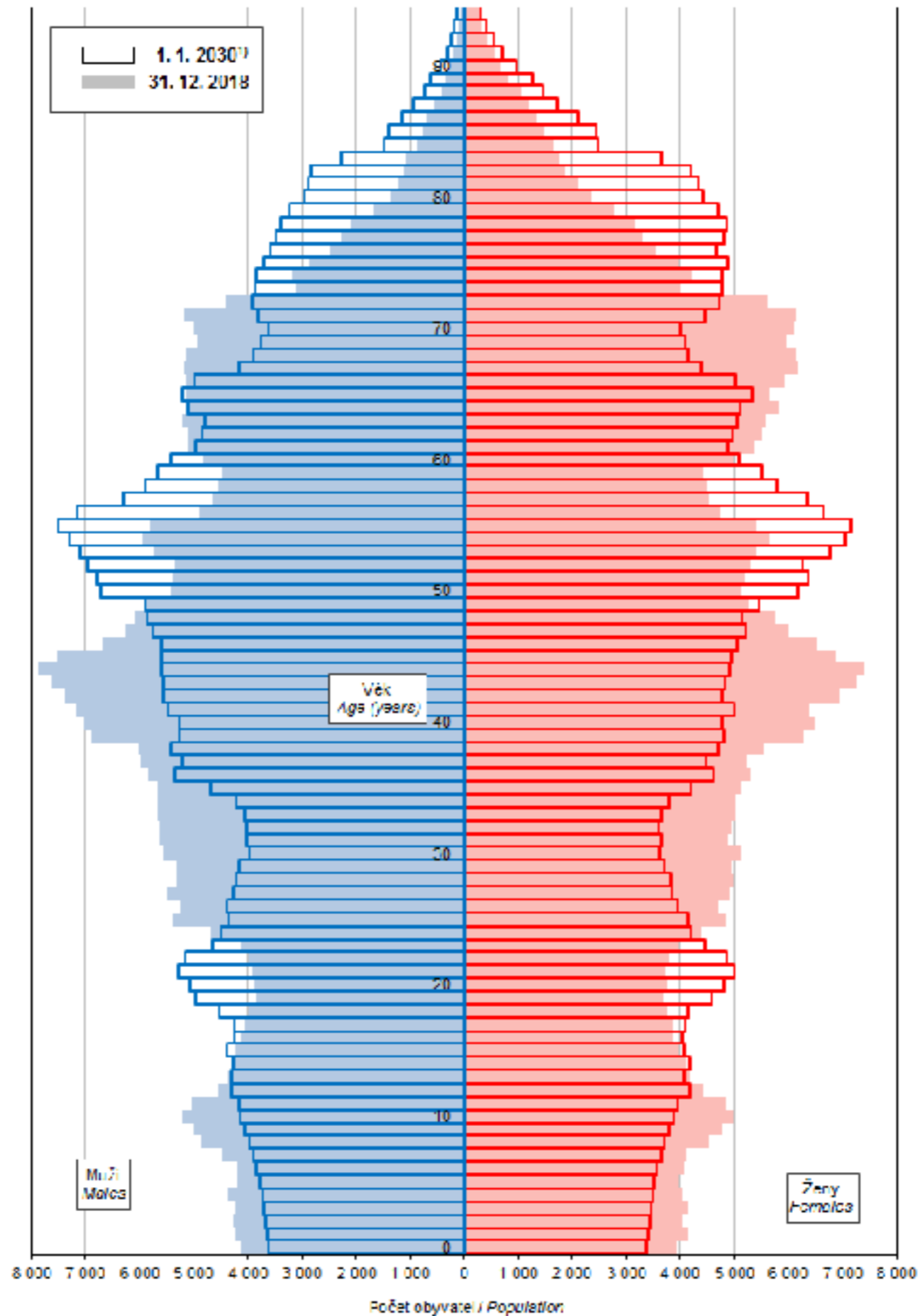
# Houslový graf (Violin plot)



# Houslový graf (Violin plot)



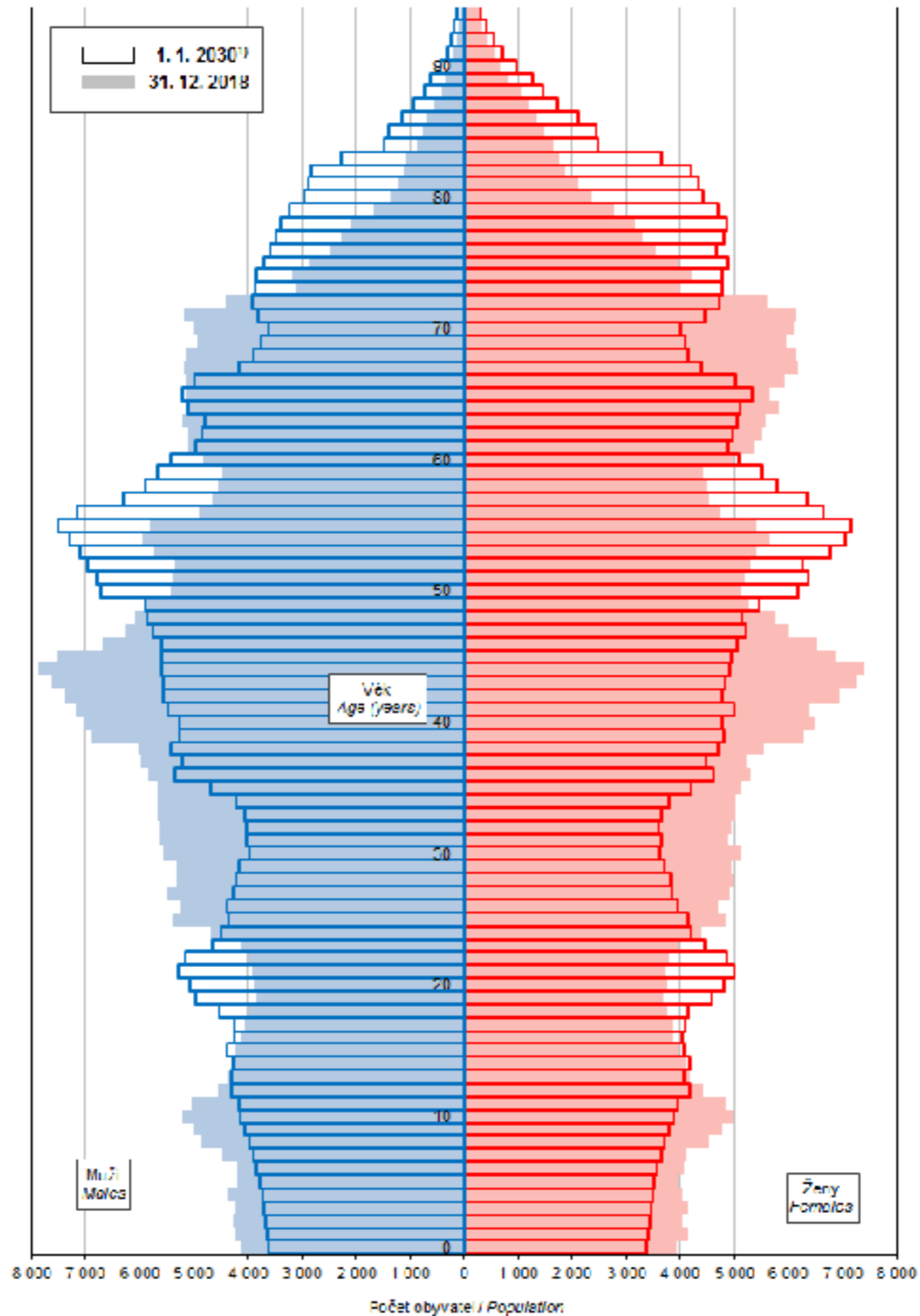
Věkové složení obyvatelstva Ústeckého kraje k 31. 12. 2018 a k 1. 1. 2030  
 Age distribution of the population in the Ústecký Region as at 31 December 2018 and 1 January 2030



<sup>1</sup>Zdroj: Projektce obyvatelstva v krajích ČR do roku 2070

<sup>2</sup>Source: ÚZSÚ publication "Projektce obyvatelstva v krajích ČR do roku 2070" (Úzoch only)

Věkové složení obyvatelstva Ústeckého kraje k 31. 12. 2018 a k 1. 1. 2030  
 Age distribution of the population in the Ústecký Region as at 31 December 2018 and 1 January 2030

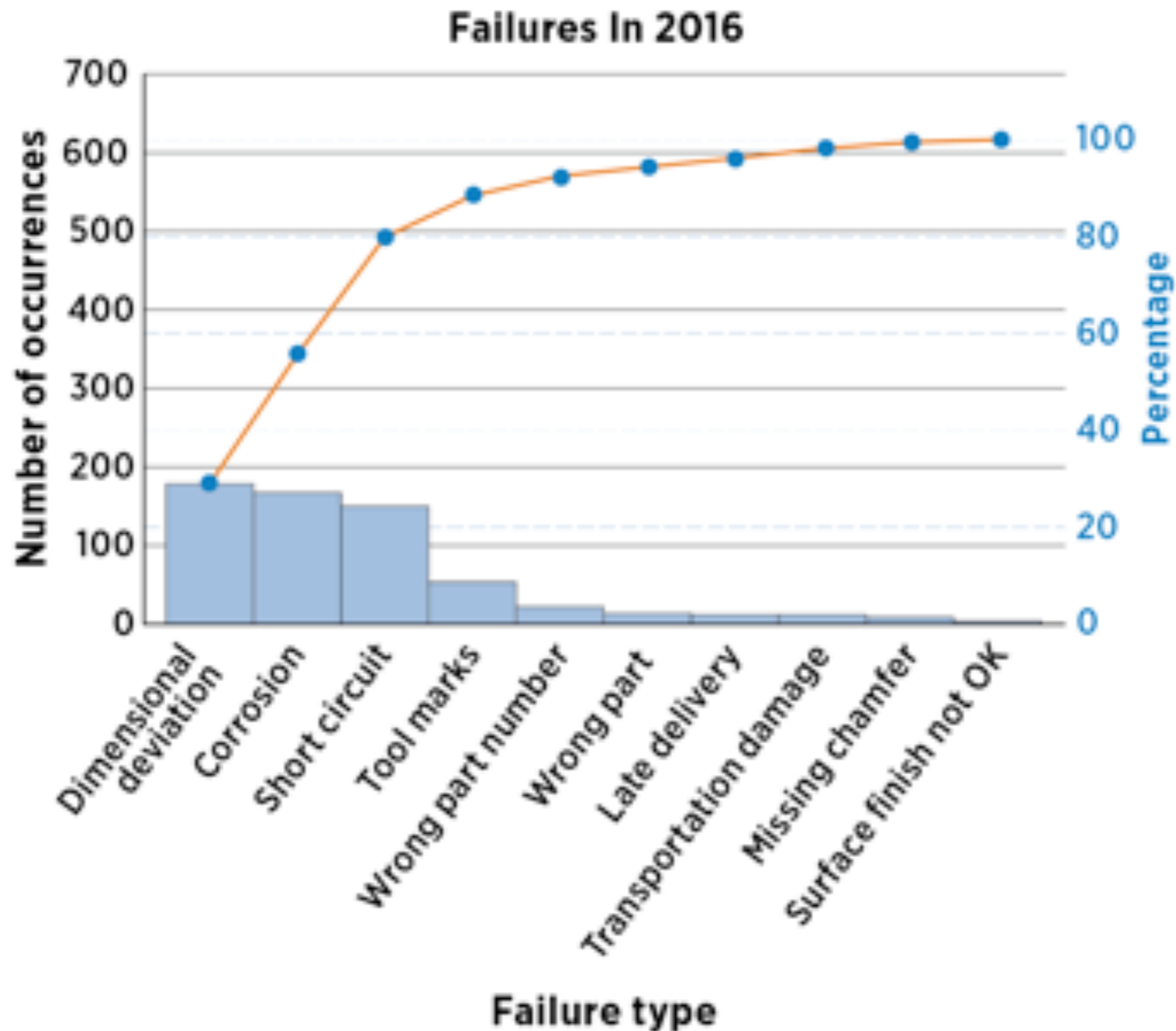


<sup>1</sup>Zdroj: Projektce obyvatelstva v krajích ČR do roku 2070

<sup>2</sup>Source: UZSO publication "Projektce obyvatelstva v krajích ČR do roku 2070" (Czech only)

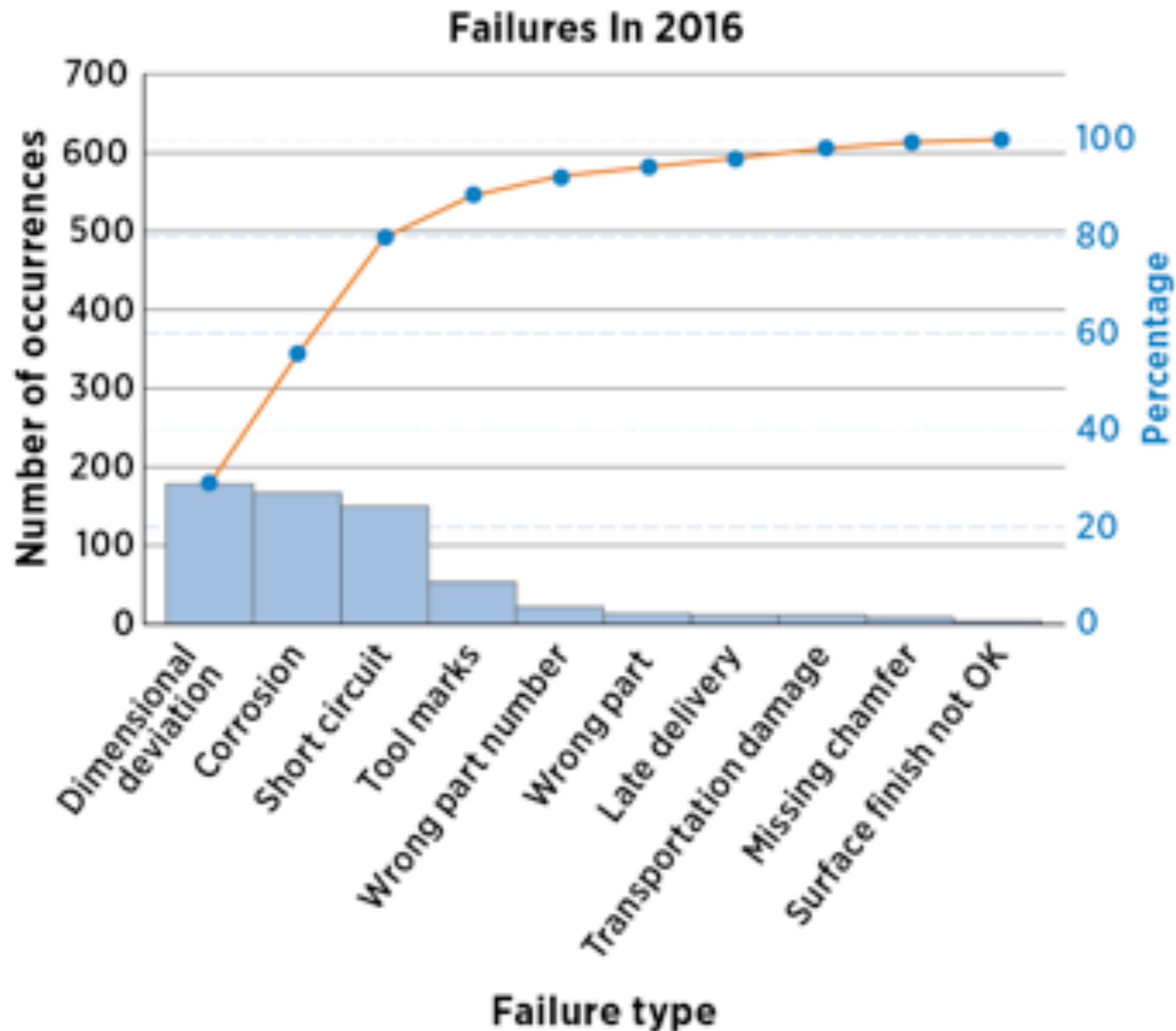


# Paretův graf





# Paretův graf



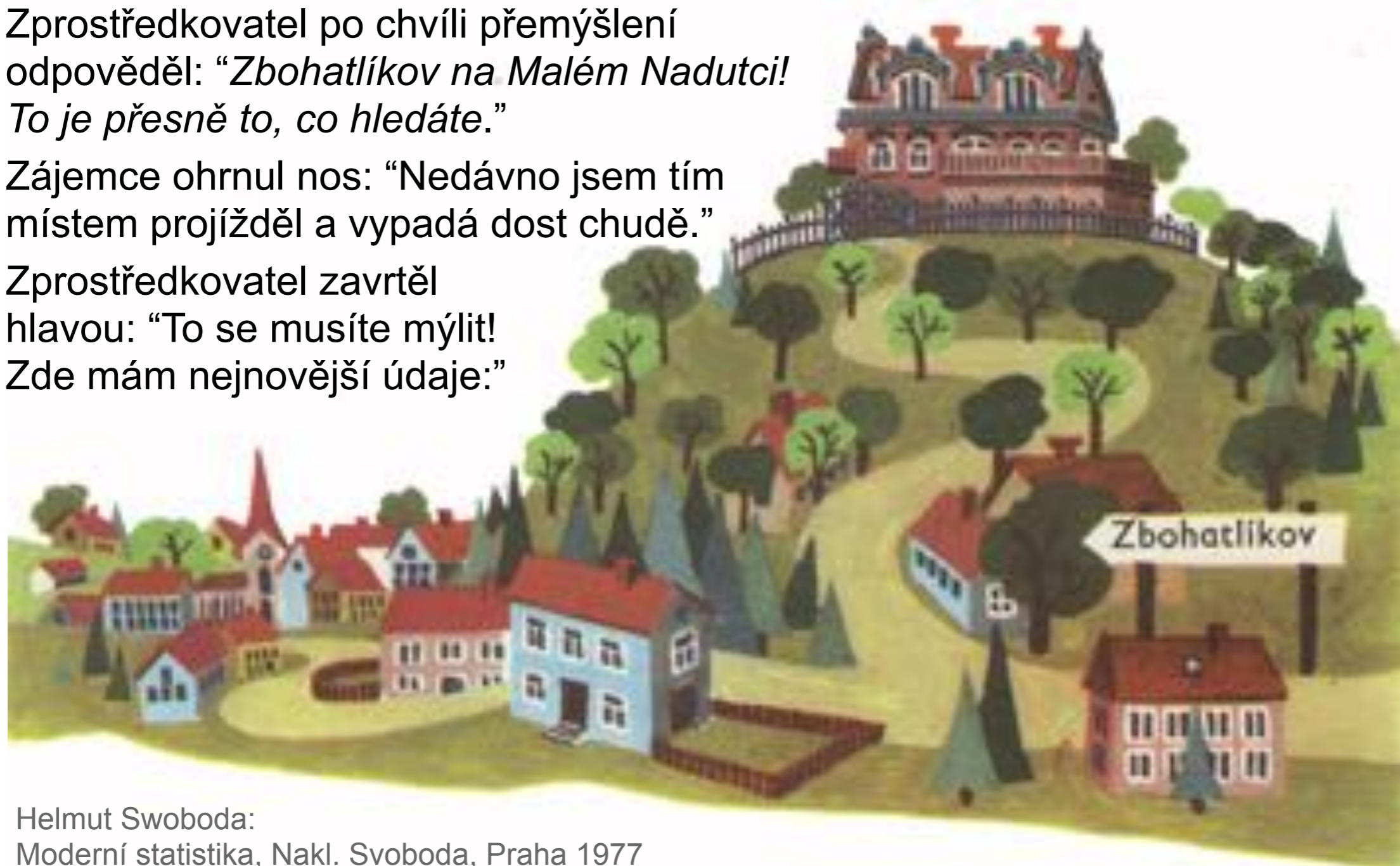
# Pohádka o Zbohatlíkově

V jedné malé rozvinuté zemi, na kraji Evropské unie, přišel mladý podnikatel do realitní kanceláře a řekl: *“Chtěl bych pozemek na venkově, s lesem, loukami, ne příliš daleko od města, v pěkné krajině, za kterou by se člověk nemusel stydět. Samozřejmě že cenově výhodný.”*

Zprostředkovatel po chvíli přemýšlení odpověděl: *“Zbohatlíkov na Malém Nadutci! To je přesně to, co hledáte.”*

Zájemce ohrnul nos: *“Nedávno jsem tím místem projížděl a vypadá dost chudě.”*

Zprostředkovatel zavrtěl hlavou: *“To se musíte mýlit! Zde mám nejnovější údaje:”*





# Pohádka o Zbohatlíkově

V jedné malé rozvinuté zemi, na kraji Evropské unie, přišel mladý podnikatel do realitní kanceláře a řekl: *“Chtěl bych pozemek na venkově, s lesem, loukami, ne příliš daleko od města, v pěkné krajině, za kterou by se člověk nemusel stydět. Samozřejmě že cenově výhodný.”*

Zprostředkovatel po chvíli přemýšlení odpověděl: *“Zbohatlíkov na Malém Nadutci! To je přesně to, co hledáte.”*

Zájemce ohrnul nos: *“Nedávno jsem tím místem projížděl a vypadá dost chudě.”*

Zprostředkovatel zavrtěl hlavou: *“To se musíte mýlit! Zde mám nejnovější údaje:”*



Helmut Swoboda:  
Moderní statistika, Nakl. Svoboda, Praha 1977



# Pohádka o Zbohatlíkově

- Zprostředkovatel tvrdí:



# Pohádka o Zbohatlíkově

- Zprostředkovatel tvrdí:

***Průměrný roční příjem ve Zbohatlíkově činí 82.320 tolarů.***





# Pohádka o Zbohatlíkově

- Zprostředkovatel tvrdí:

***Průměrný roční příjem ve Zbohatlíkově činí 82.320 dolarů.***

- Kupec zašel za známým ředitelem banky:

*roční příjem více než poloviny obyvatel je **29.000 dolarů a více.***



# Pohádka o Zbohatlíkově

- Zprostředkovatel tvrdí:

***Průměrný roční příjem ve Zbohatlíkově činí 82.320 dolarů.***

- Kupec zašel za známým ředitelem banky:

*roční příjem více než poloviny obyvatel je **29.000 dolarů** a více.*

- To je podivné! Co řekne okresní úřad?:

*Dostí chudé místo, průměrný příjem je kolem **29.000 dolarů**.*



# Pohádka o Zbohatlíkově

- Zprostředkovatel tvrdí:

***Průměrný roční příjem ve Zbohatlíkově činí 82.320 tolarů.***

- Kupec zašel za známým ředitelem banky:

*roční příjem více než poloviny obyvatel je **29.000 tolarů** a více.*

- To je podivné! Co řekne okresní úřad?:

*Dosti chudé místo, průměrný příjem je kolem **29.000 tolarů**.*

- Vrátil se k řediteli banky pro nové informace:

*Nejsilněji zastoupená příjmová kategorie je od **12.000** do **24.000 tolarů***

*Nejčastější příjem je poměrně přesně **18.000 tolarů**.*





# Pohádka o Zbohatlíkově

- Zprostředkovatel tvrdí:

*Průměrný roční příjem ve Zbohatlíkově činí **82.320** tolarů.*

- Kupec zašel za známým ředitelem banky:

*roční příjem více než poloviny obyvatel je **29.000** tolarů a více.*

- To je podivné! Co řekne okresní úřad?:

*Dosti chudé místo, průměrný příjem je kolem **29.000** tolarů.*

- Vrátil se k řediteli banky pro nové informace:

*Nejsilněji zastoupená příjmová kategorie je od **12.000** do **24.000** tolarů*

*Nejčastější příjem je poměrně přesně **18.000** tolarů.*

- Rozhněvaný kupec jede za učitelem Počtářem, kam ho poslali. Ten tvrdí, že situace je neutěšená:

*Dvě třetiny rodin mají méně než **30.000** tolarů.*



# Pohádka o Zbohatlíkově

- Zprostředkovatel tvrdí:

***Průměrný roční příjem ve Zbohatlíkově činí 82.320 tolarů.***

- Kupec zašel za známým ředitelem banky:

*roční příjem více než poloviny obyvatel je **29.000 tolarů** a více.*

- To je podivné! Co řekne okresní úřad?:

*Dosti chudé místo, průměrný příjem je kolem **29.000 tolarů**.*

- Vrátil se k řediteli banky pro nové informace:

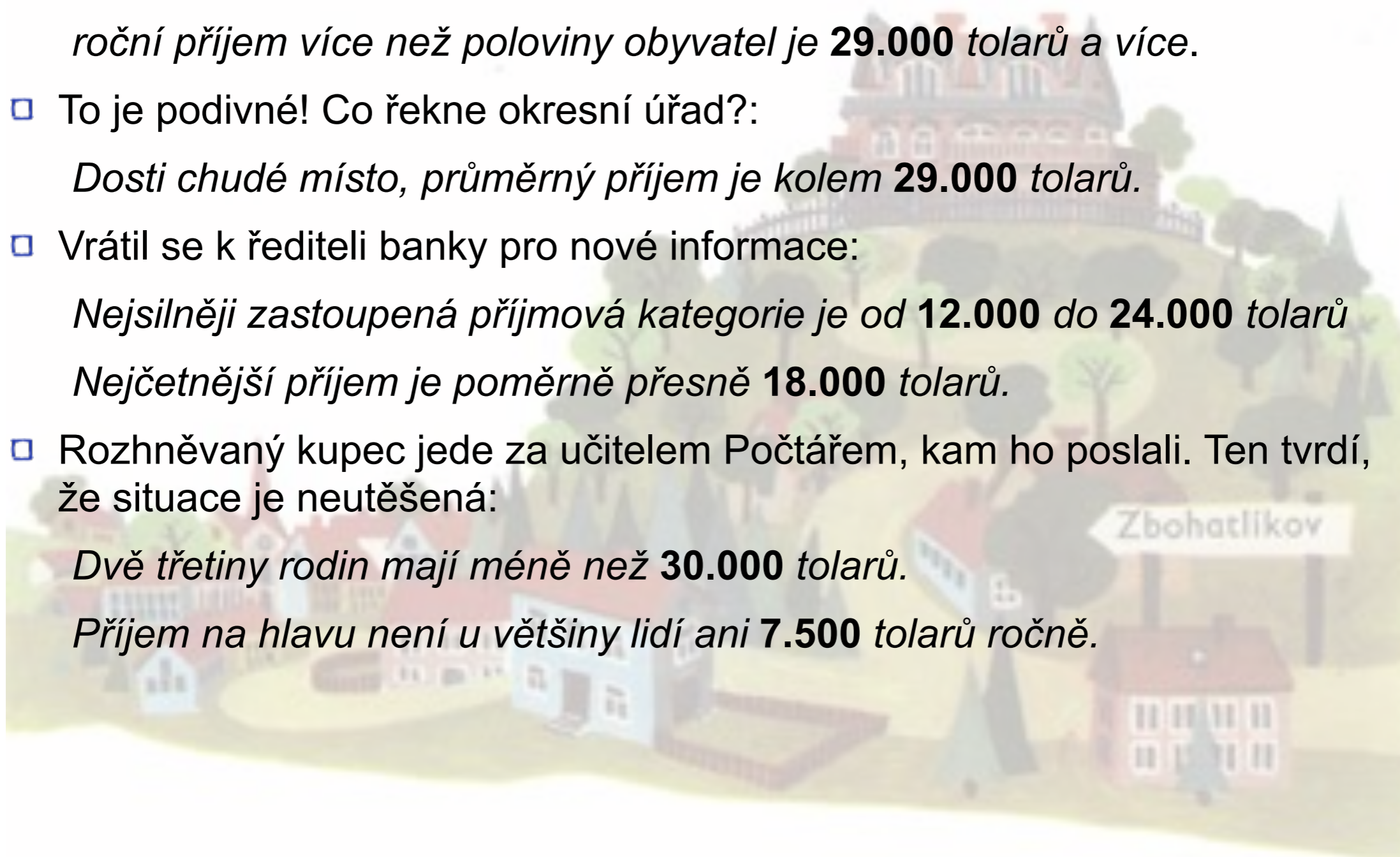
*Nejsilněji zastoupená příjmová kategorie je od **12.000** do **24.000 tolarů***

*Nejčastější příjem je poměrně přesně **18.000 tolarů**.*

- Rozhněvaný kupec jede za učitelem Počtářem, kam ho poslali. Ten tvrdí, že situace je neutěšená:

*Dvě třetiny rodin mají méně než **30.000 tolarů**.*

*Příjem na hlavu není u většiny lidí ani **7.500 tolarů** ročně.*





# Pohádka o Zbohatlíkově

- Zprostředkovatel tvrdí:

***Průměrný roční příjem ve Zbohatlíkově činí 82.320 tolarů.***

- Kupec zašel za známým ředitelem banky:

*roční příjem více než poloviny obyvatel je **29.000 tolarů** a více.*

- To je podivné! Co řekne okresní úřad?:

*Dosti chudé místo, průměrný příjem je kolem **29.000 tolarů**.*

- Vrátil se k řediteli banky pro nové informace:

*Nejsilněji zastoupená příjmová kategorie je od **12.000** do **24.000 tolarů***

*Nejčastější příjem je poměrně přesně **18.000 tolarů**.*

- Rozhněvaný kupec jede za učitelem Počtářem, kam ho poslali. Ten tvrdí, že situace je neutěšená:

*Dvě třetiny rodin mají méně než **30.000 tolarů**.*

*Příjem na hlavu není u většiny lidí ani **7.500 tolarů** ročně.*

*80% obyvatel má ročně méně než **25.000 tolarů***

# Pohádka o Zbohatlíkově

- Zprostředkovatel tvrdí:

*Průměrný roční příjem ve Zbohatlíkově činí **82.320** tolarů.*

- Kupec zašel za známým ředitelem banky:

*roční příjem více než poloviny obyvatel je **29.000** tolarů a více.*

- To je podivné! Co řekne okresní úřad?:

*Dosti chudé místo, průměrný příjem je kolem **29.000** tolarů.*

- Vrátil se k řediteli banky pro nové informace:

*Nejsilněji zastoupená příjmová kategorie je od **12.000** do **24.000** tolarů*

*Nejčastější příjem je poměrně přesně **18.000** tolarů.*

- Rozhněvaný kupec jede za učitelem Počtářem, kam ho poslali. Ten tvrdí, že situace je neutěšená:

*Dvě třetiny rodin mají méně než **30.000** tolarů.*

*Příjem na hlavu není u většiny lidí ani **7.500** tolarů ročně.*

*80% obyvatel má ročně méně než **25.000** tolarů*

**Kdo z nich lže?**



# Pohádka o Zbohatlíkově

- Zprostředkovatel tvrdí:

*Průměrný roční příjem ve Zbohatlíkově činí 82.320 tolarů.*

- Kupec zašel za známým ředitelem banky:

*roční příjem více než poloviny obyvatel je 29.000 tolarů a více.*

- To je podivné! Co řekne okresní úřad?:

*Dosti chudé místo, průměrný příjem je kolem 29.000 tolarů.*

- Vrátil se k řediteli banky pro nové informace:

*Nejsilněji zastoupená příjmová kategorie je od 12.000 do 24.000 tolarů*

*Nejčastější příjem je poměrně přesně 18.000 tolarů.*

- Rozhněvaný kupec jede za učitelem Počtářem, kam ho poslali. Ten tvrdí, že situace je neutěšená:

*Dvě třetiny rodin mají méně než 30.000 tolarů.*

*Příjem na hlavu není u většiny lidí ani 7.500 tolarů ročně.*

*80% obyvatel má ročně méně než 25.000 tolarů*

**Kdo z nich lže?**



# Pohádka o Zbohatlíkově

Údaje o ročním příjmu 25 rodin ze Zbohatlíkova, n je počet členů domácnosti:

| roční příjem | n | roční příjem | n | roční příjem | n | roční příjem | n |
|--------------|---|--------------|---|--------------|---|--------------|---|
| 1,200.000    | 3 | 60.000       | 1 | 45.000       | 2 | 29.000       | 3 |
| 150.000      | 5 | 51.000       | 3 | 42.000       | 2 | 26.000       | 4 |
| 86.000       | 4 | 49.000       | 4 | 38.000       | 4 | 24.000       | 4 |
| 37.000       | 3 | 20.000       | 7 | 14.000       | 1 | 18.000       | 4 |
| 35.000       | 5 | 18.000       | 3 | 13.000       | 4 | 16.000       | 3 |
| 32.000       | 3 | 18.000       | 8 | 11.000       | 1 | 16.000       | 2 |
|              |   |              |   |              |   | 10.000       | 2 |







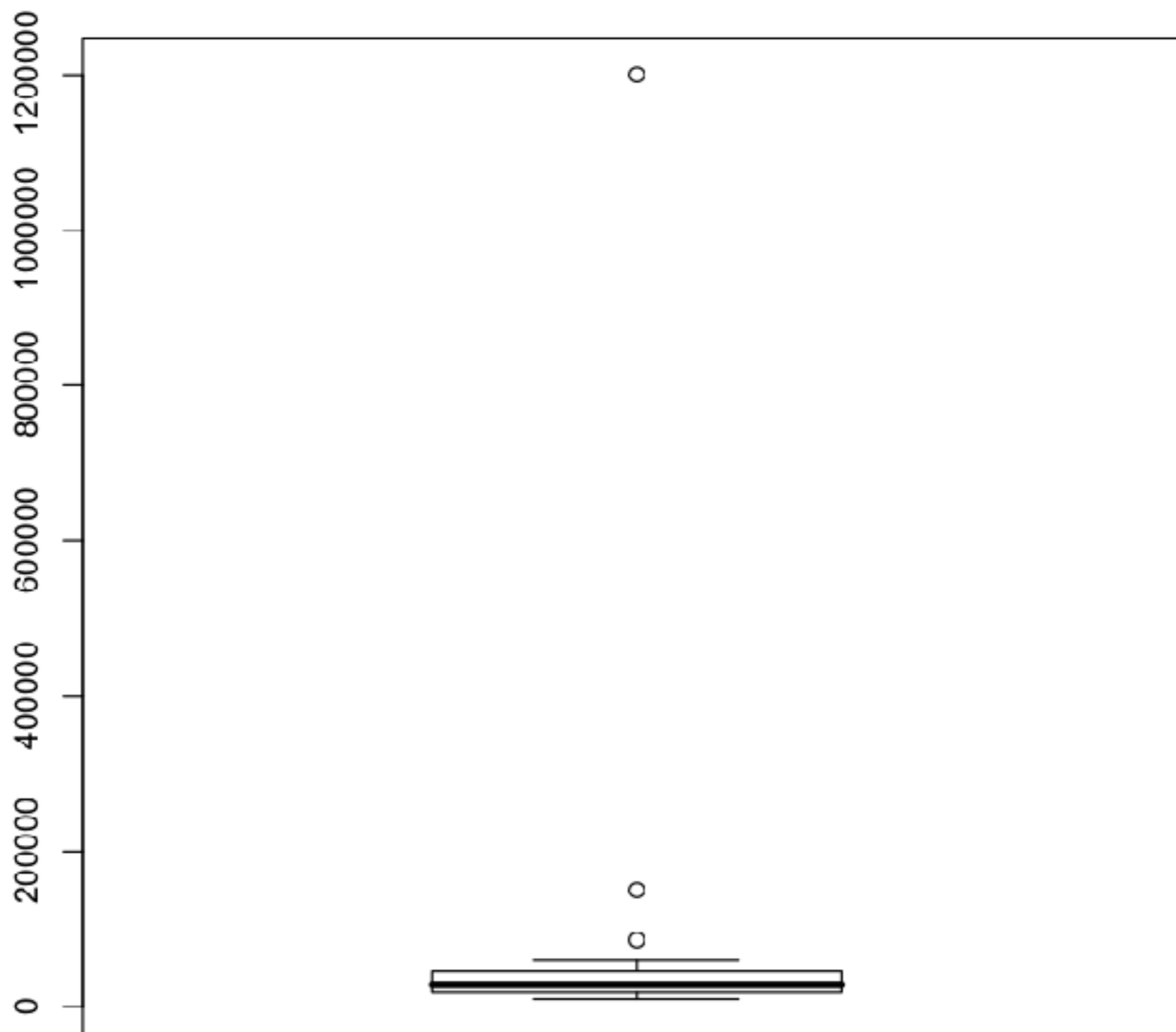






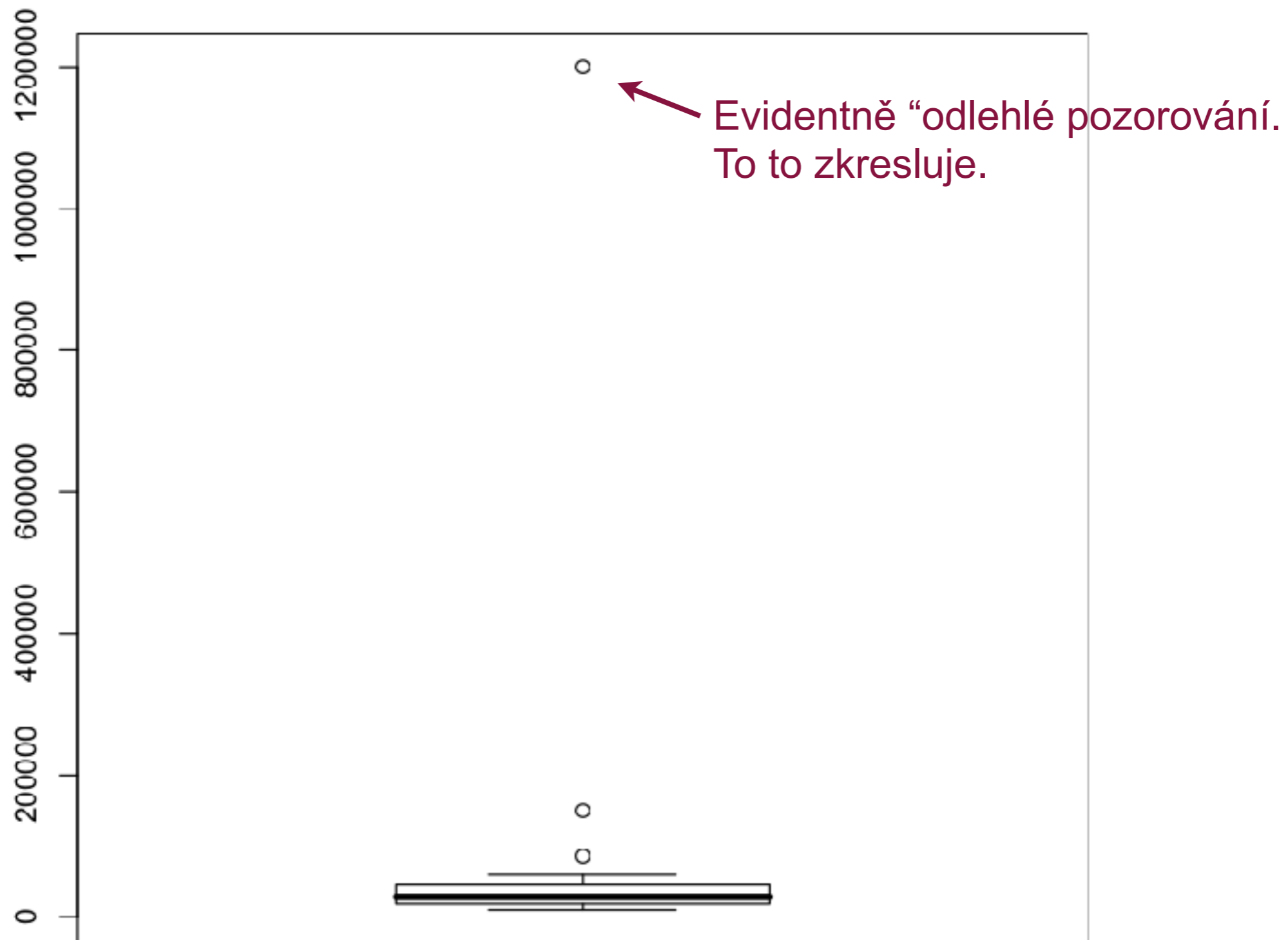
# Pohádka o Zbohatlívě

... a co “Box&Whiskers” diagram?



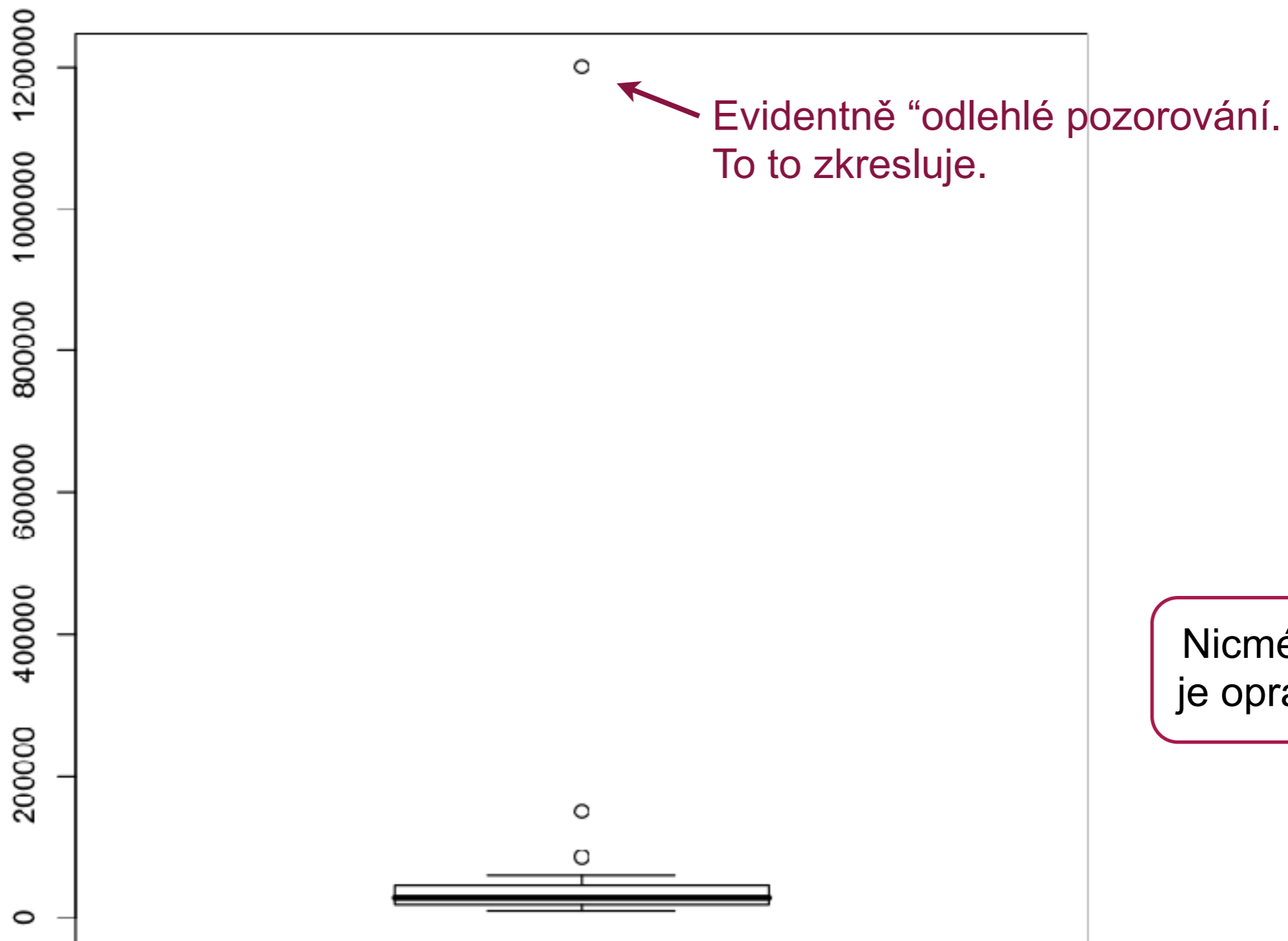
# Pohádka o Zbohatlívě

... a co “Box&Whiskers” diagram?



# Pohádka o Zbohatlívě

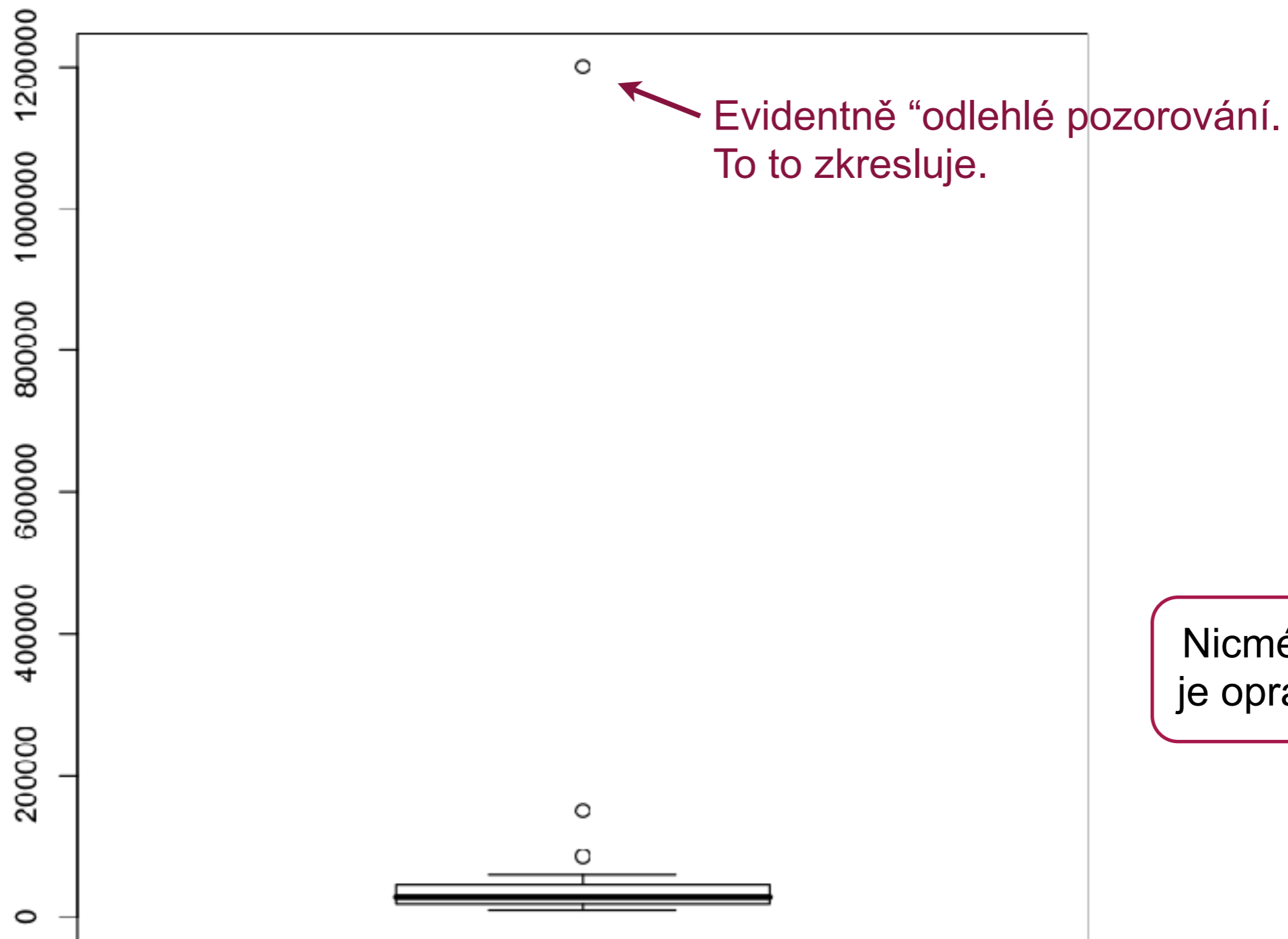
... a co “Box&Whiskers” diagram?



Nicméně, medián  
je opravdu 29.000

# Pohádka o Zbohatlíkově

... a co “Box&Whiskers” diagram?

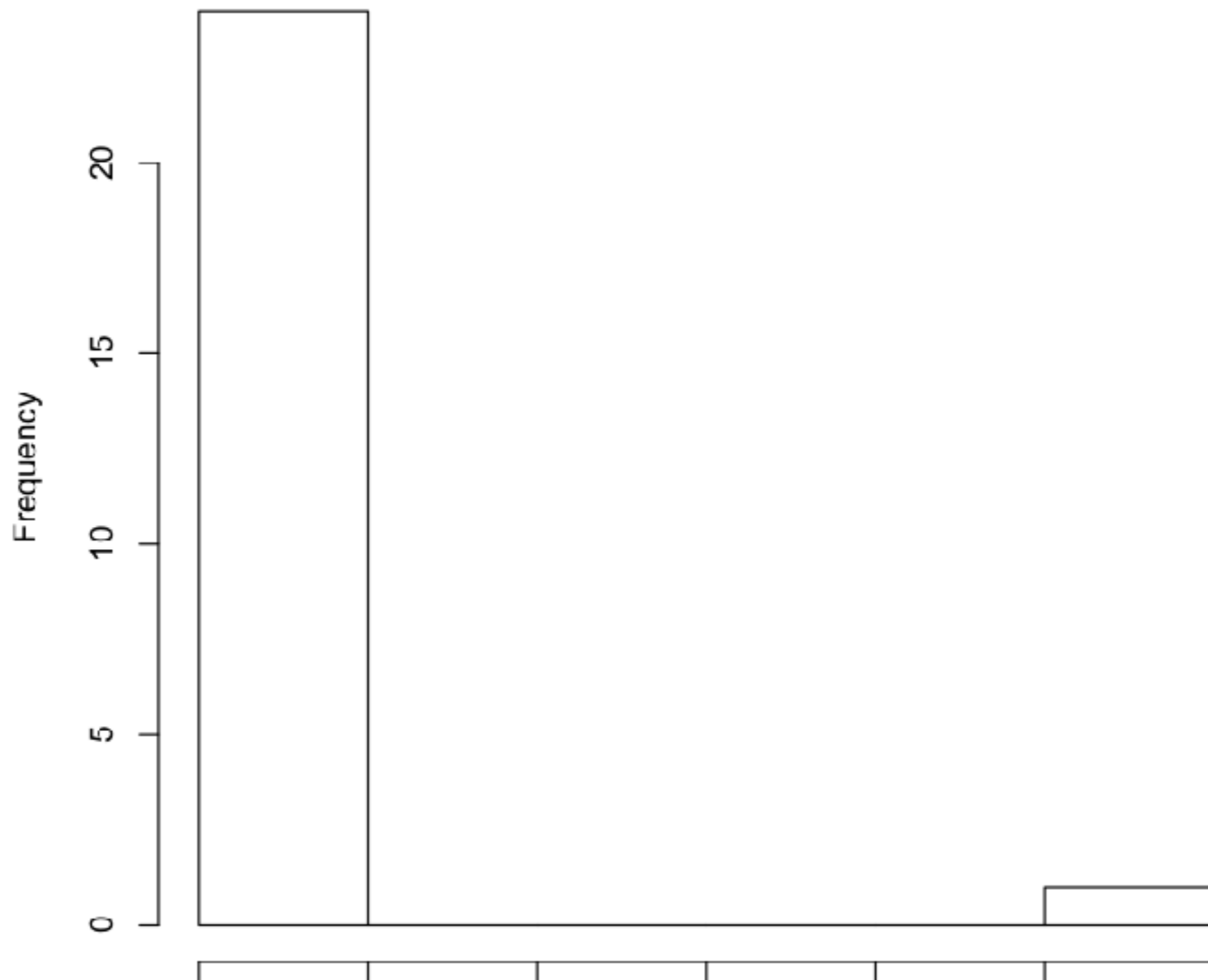


Nicméně, medián je opravdu 29.000



# Pohádka o Zbohatlíkově

stejně dopadne i histogram:



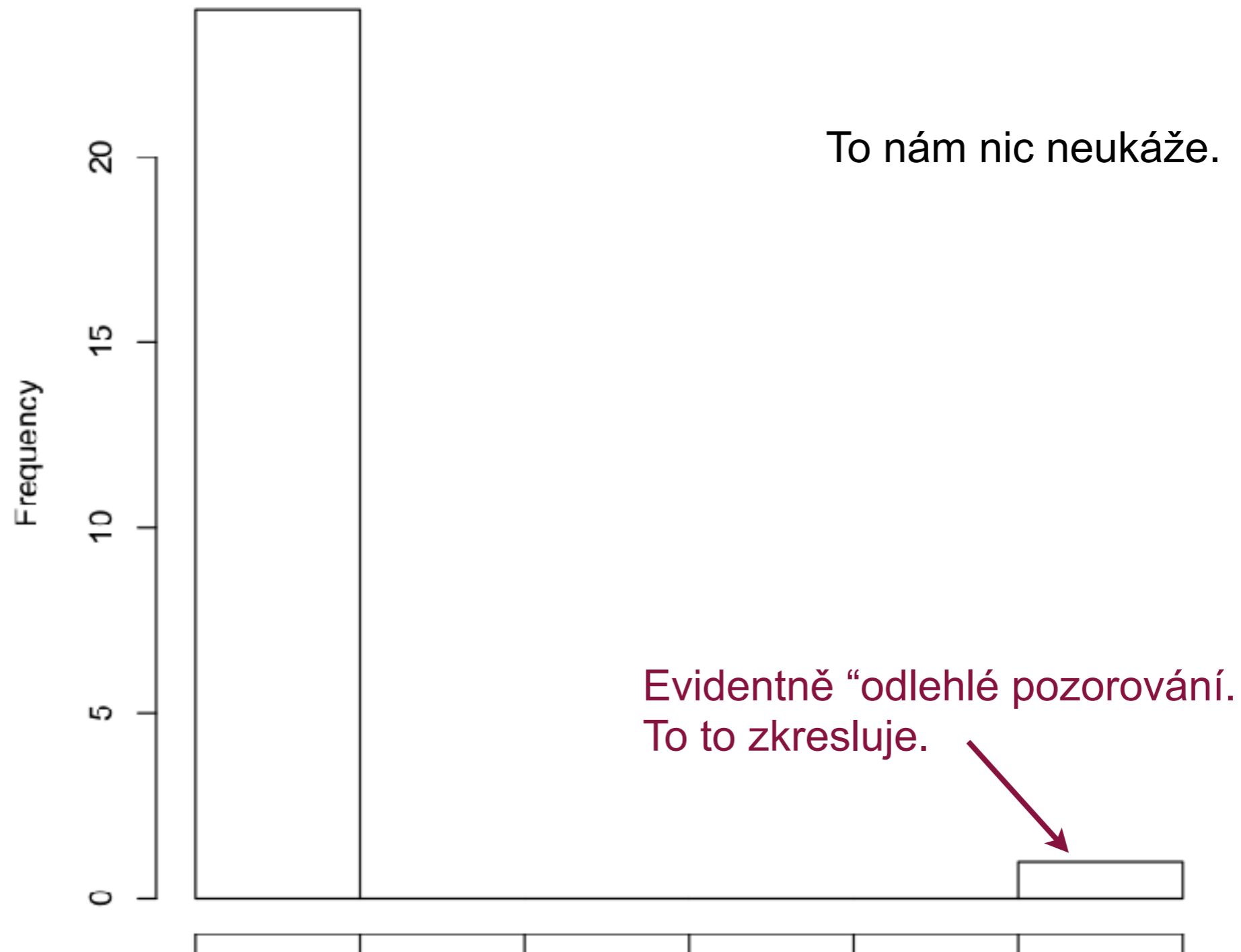
# Pohádka o Zbohatlíkově

stejně dopadne i histogram:



# Pohádka o Zbohatlíkově

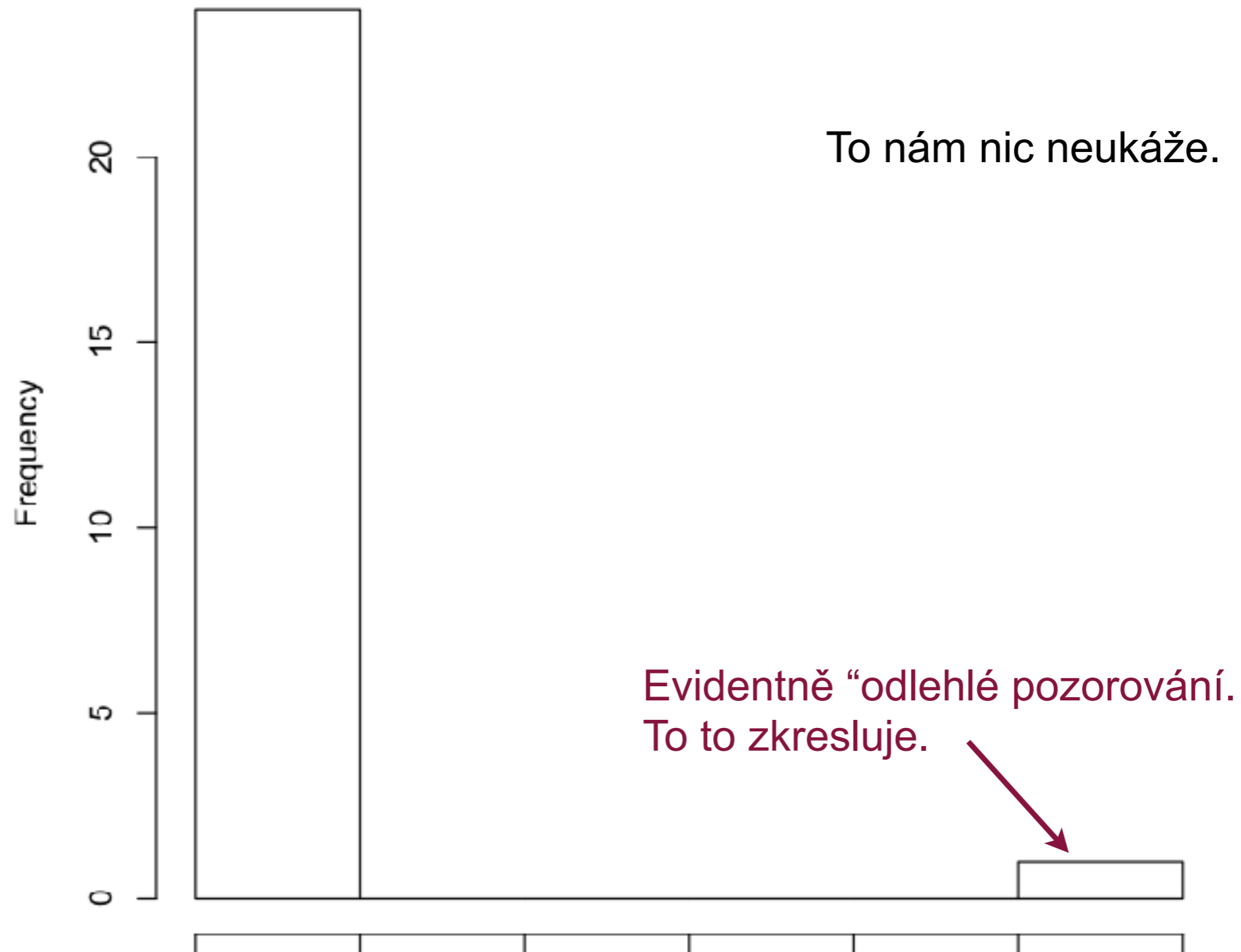
stejně dopadne i histogram:





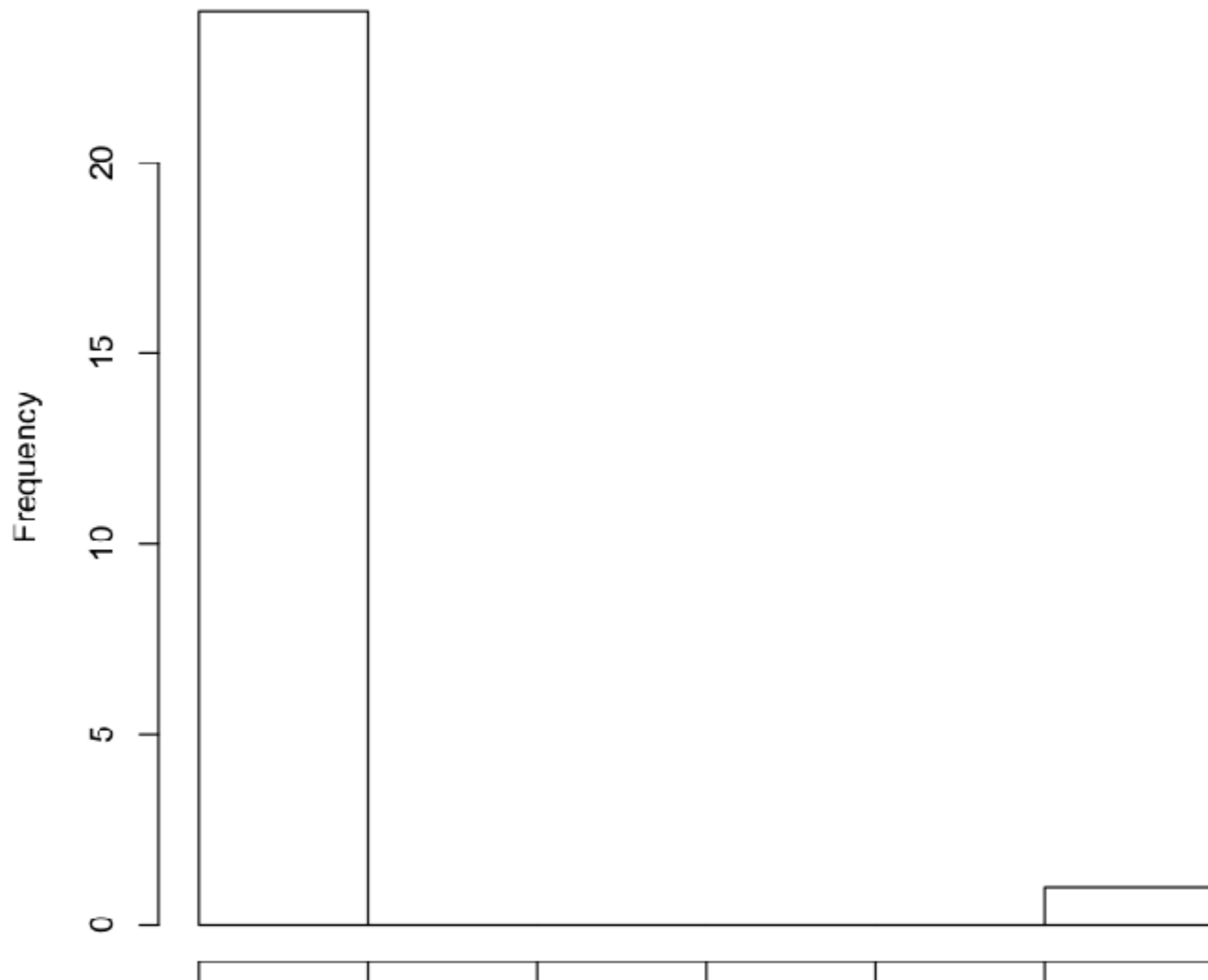
# Pohádka o Zbohatlíkově

stejně dopadne i histogram:



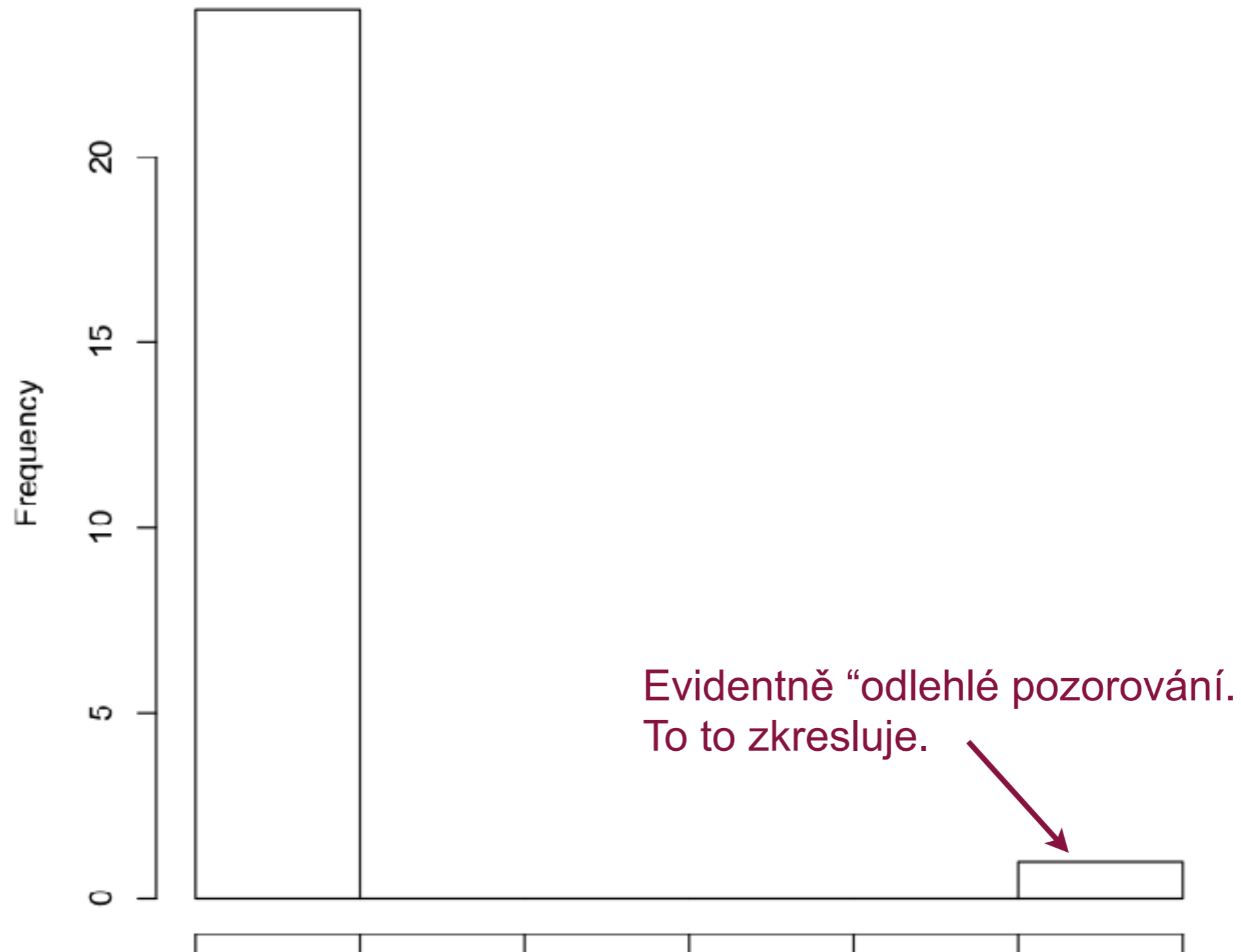
# Pohádka o Zbohatlívě

stejně dopadne i histogram:



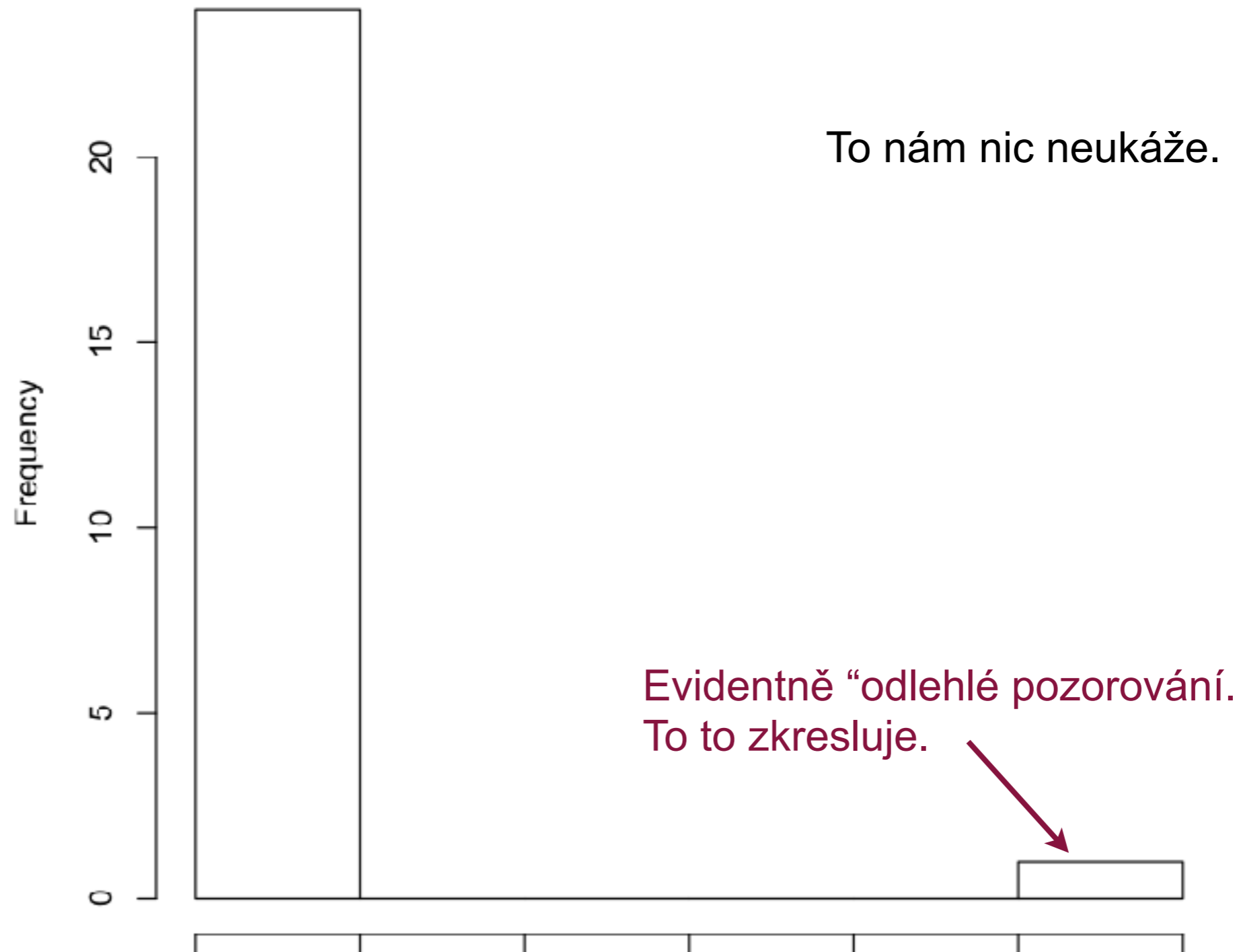
# Pohádka o Zbohatlíkově

stejně dopadne i histogram:



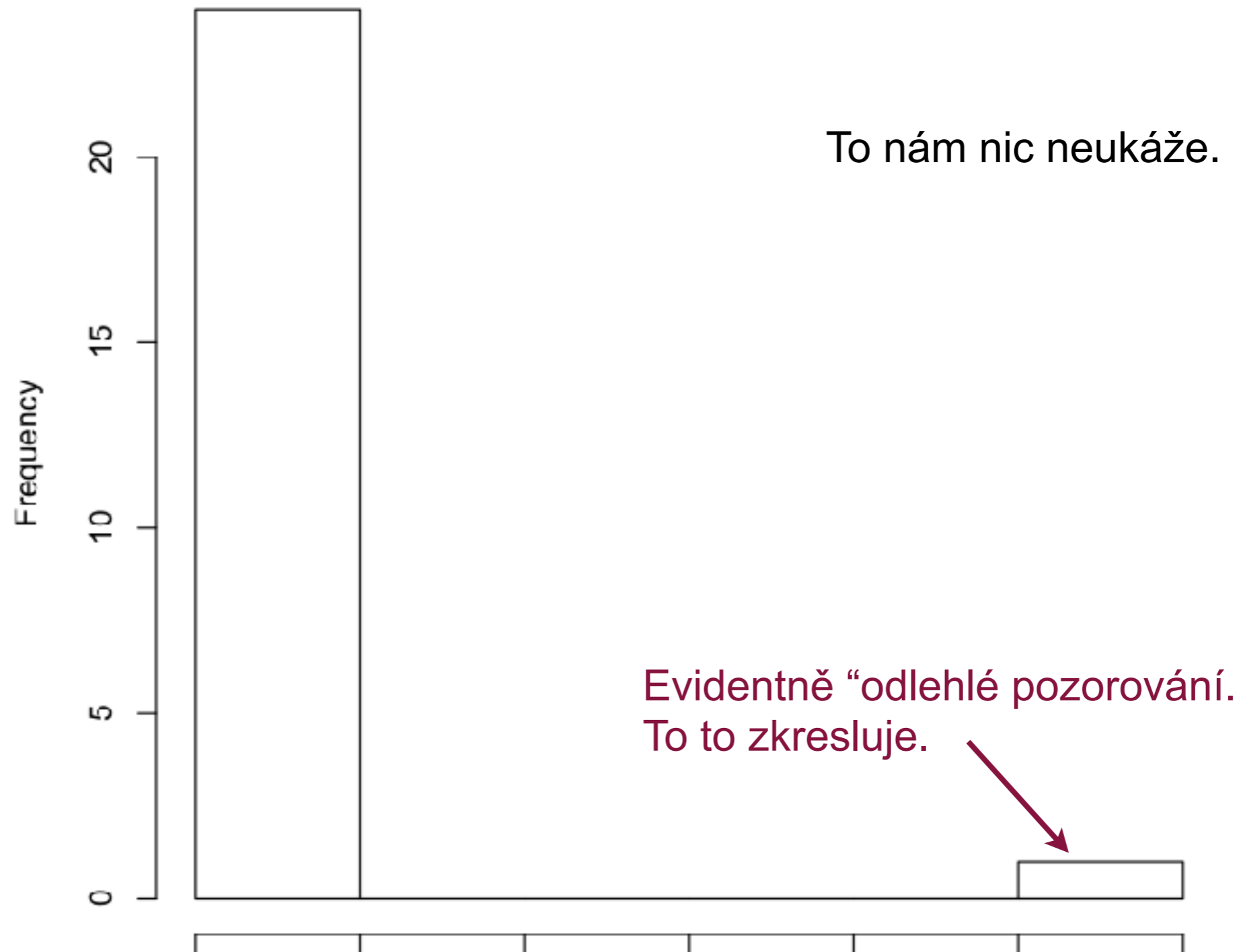
# Pohádka o Zbohatlíkově

stejně dopadne i histogram:



# Pohádka o Zbohatlíkově

stejně dopadne i histogram:



# Pohádka o Zbohatlíkově

Údaje o ročním příjmu 25 rodin ze Zbohatlíkova, n je počet členů domácnosti:

| roční příjem | n | roční příjem | n | roční příjem | n |        |   |
|--------------|---|--------------|---|--------------|---|--------|---|
| 1,200.000    | 3 | 60.000       | 1 | 45.000       | 2 | 29.000 | 3 |
| 150.000      | 5 | 51.000       | 3 | 42.000       | 2 | 26.000 | 4 |
| 86.000       | 4 | 49.000       | 4 | 38.000       | 4 | 24.000 | 4 |
| 37.000       | 3 | 20.000       | 7 | 14.000       | 1 | 18.000 | 4 |
| 35.000       | 5 | 18.000       | 3 | 13.000       | 4 | 16.000 | 3 |
| 32.000       | 3 | 18.000       | 8 | 11.000       | 1 | 16.000 | 2 |
|              |   |              |   |              |   | 10.000 | 2 |

Odstraníme na chvíli extrémní (odlehlou) hodnotu :



# Pohádka o Zbohatlíkově

Údaje o ročním příjmu 25 rodin ze Zbohatlíkova, n je počet členů domácnosti:

| roční příjem | n | roční příjem | n | roční příjem | n |        |   |
|--------------|---|--------------|---|--------------|---|--------|---|
| 1,200.000    | 3 | 60.000       | 1 | 45.000       | 2 | 29.000 | 3 |
| 150.000      | 5 | 51.000       | 3 | 42.000       | 2 | 26.000 | 4 |
| 86.000       | 4 | 49.000       | 4 | 38.000       | 4 | 24.000 | 4 |
| 37.000       | 3 | 20.000       | 7 | 14.000       | 1 | 18.000 | 4 |
| 35.000       | 5 | 18.000       | 3 | 13.000       | 4 | 16.000 | 3 |
| 32.000       | 3 | 18.000       | 8 | 11.000       | 1 | 16.000 | 2 |
|              |   |              |   |              |   | 10.000 | 2 |

Odstraníme na chvíli extrémní (odlehlou) hodnotu :

0 | 1111222222223334444

0 | 55569

1 |

1 | 5





# Pohádka o Zbohatlíkově

Údaje o ročním příjmu 25 rodin ze Zbohatlíkova, n je počet členů domácnosti:

| roční příjem | n | roční příjem | n | roční příjem | n |        |   |
|--------------|---|--------------|---|--------------|---|--------|---|
| 1,200.000    | 3 | 60.000       | 1 | 45.000       | 2 | 29.000 | 3 |
| 150.000      | 5 | 51.000       | 3 | 42.000       | 2 | 26.000 | 4 |
| 86.000       | 4 | 49.000       | 4 | 38.000       | 4 | 24.000 | 4 |
| 37.000       | 3 | 20.000       | 7 | 14.000       | 1 | 18.000 | 4 |
| 35.000       | 5 | 18.000       | 3 | 13.000       | 4 | 16.000 | 3 |
| 32.000       | 3 | 18.000       | 8 | 11.000       | 1 | 16.000 | 2 |
|              |   |              |   |              |   | 10.000 | 2 |

Odstraníme na chvíli extrémní (odlehlou) hodnotu :

0 | 1111222222223334444

0 | 55569

1 |

1 | 5

To už je trochu lepší!



# Pohádka o Zbohatlíkově

Údaje o ročním příjmu 25 rodin ze Zbohatlíkova, n je počet členů domácnosti:

| roční příjem | n | roční příjem | n | roční příjem | n |        |   |
|--------------|---|--------------|---|--------------|---|--------|---|
| 1,200.000    | 3 | 60.000       | 1 | 45.000       | 2 | 29.000 | 3 |
| 150.000      | 5 | 51.000       | 3 | 42.000       | 2 | 26.000 | 4 |
| 86.000       | 4 | 49.000       | 4 | 38.000       | 4 | 24.000 | 4 |
| 37.000       | 3 | 20.000       | 7 | 14.000       | 1 | 18.000 | 4 |
| 35.000       | 5 | 18.000       | 3 | 13.000       | 4 | 16.000 | 3 |
| 32.000       | 3 | 18.000       | 8 | 11.000       | 1 | 16.000 | 2 |
|              |   |              |   |              |   | 10.000 | 2 |

Odstraníme na chvíli extrémní (odlehlou) hodnotu :

0 | 1111222222223334444

0 | 55569

1 |

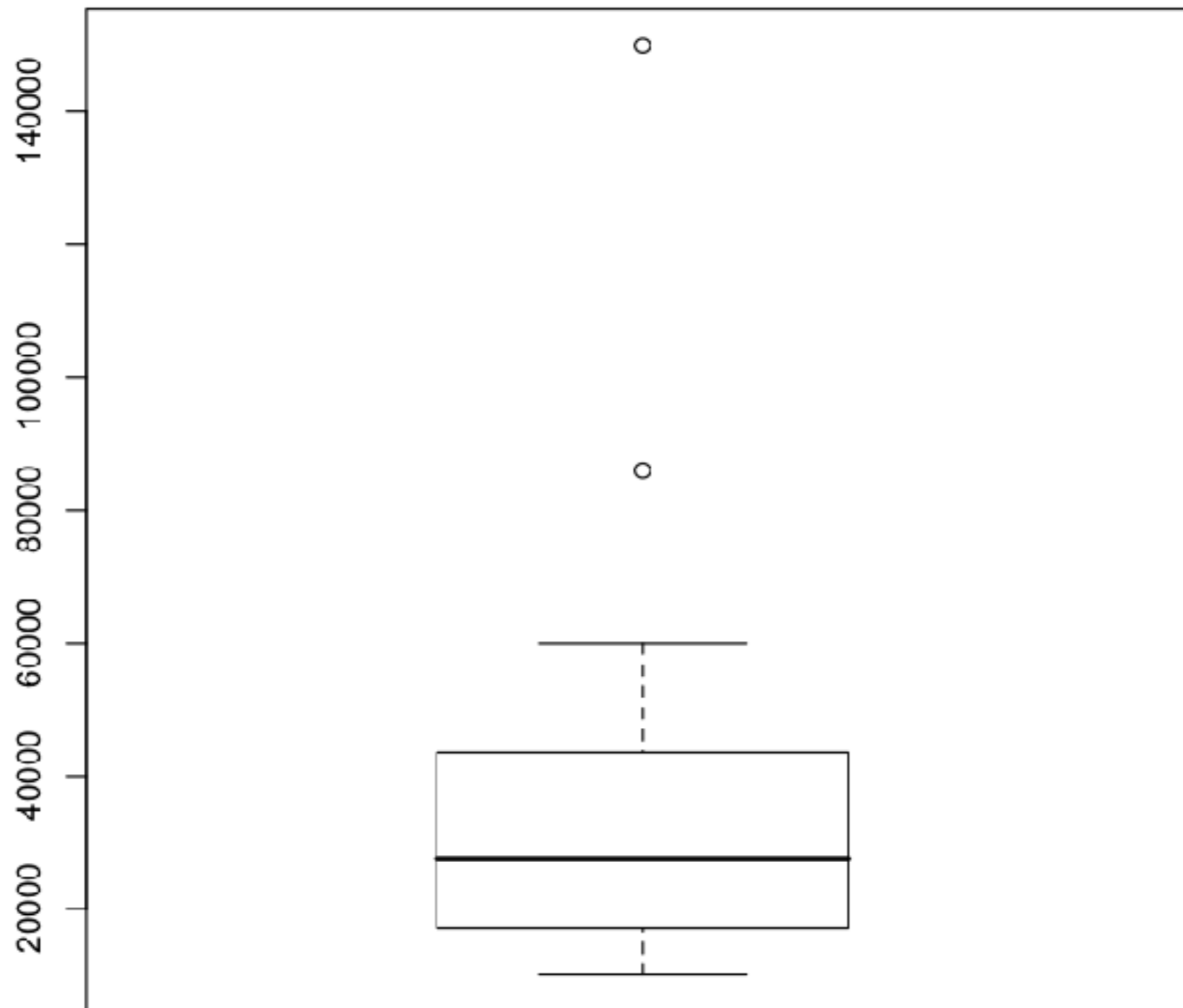
1 | 5

To už je trochu lepší!



# Pohádka o Zbohatlívě

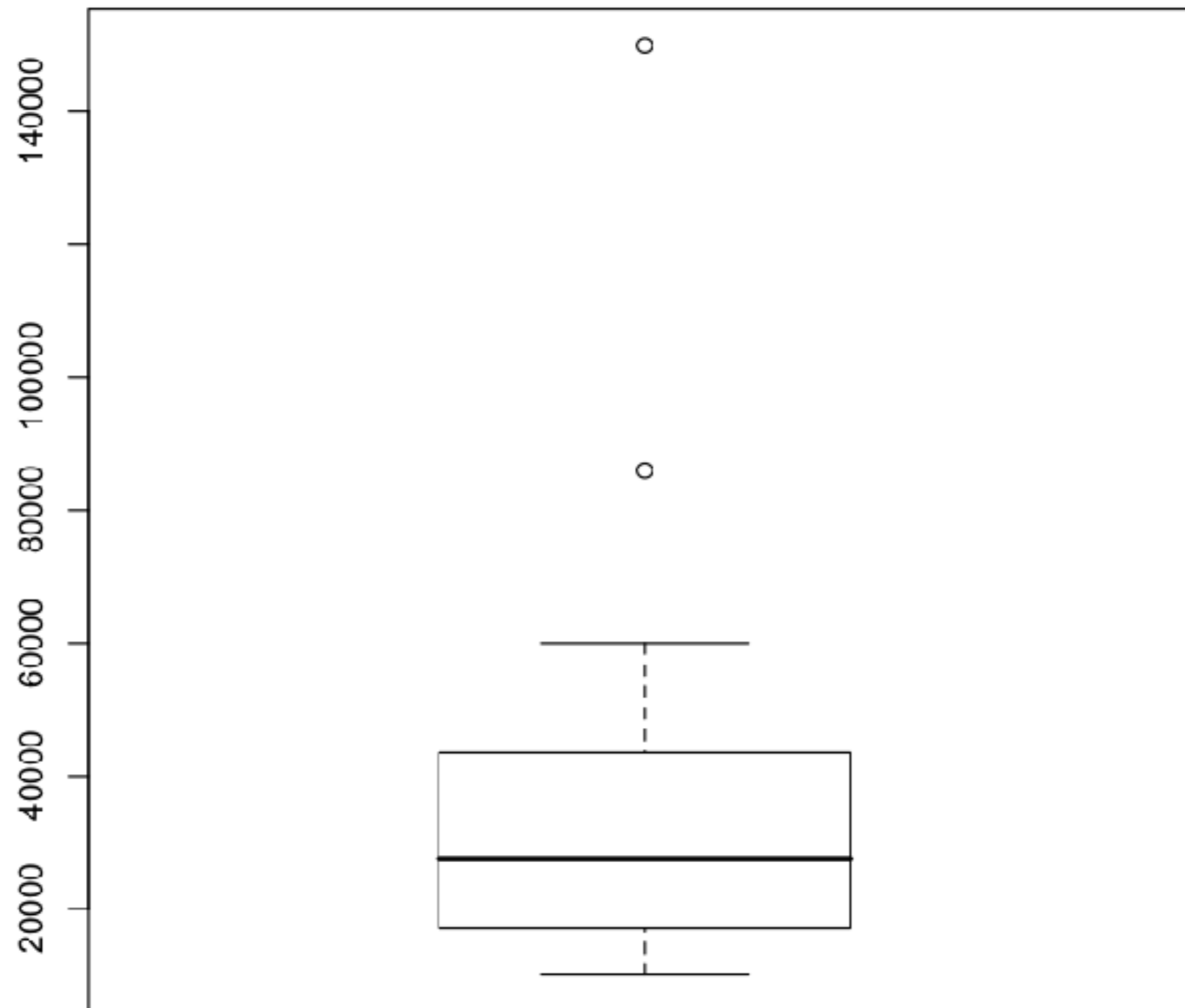
Odstraníme na chvíli extrémní (odlehlou) hodnotu :



$$X_{\text{med}} = 29.000$$

# Pohádka o Zbohatlívě

Odstraníme na chvíli extrémní (odlehlu) hodnotu :

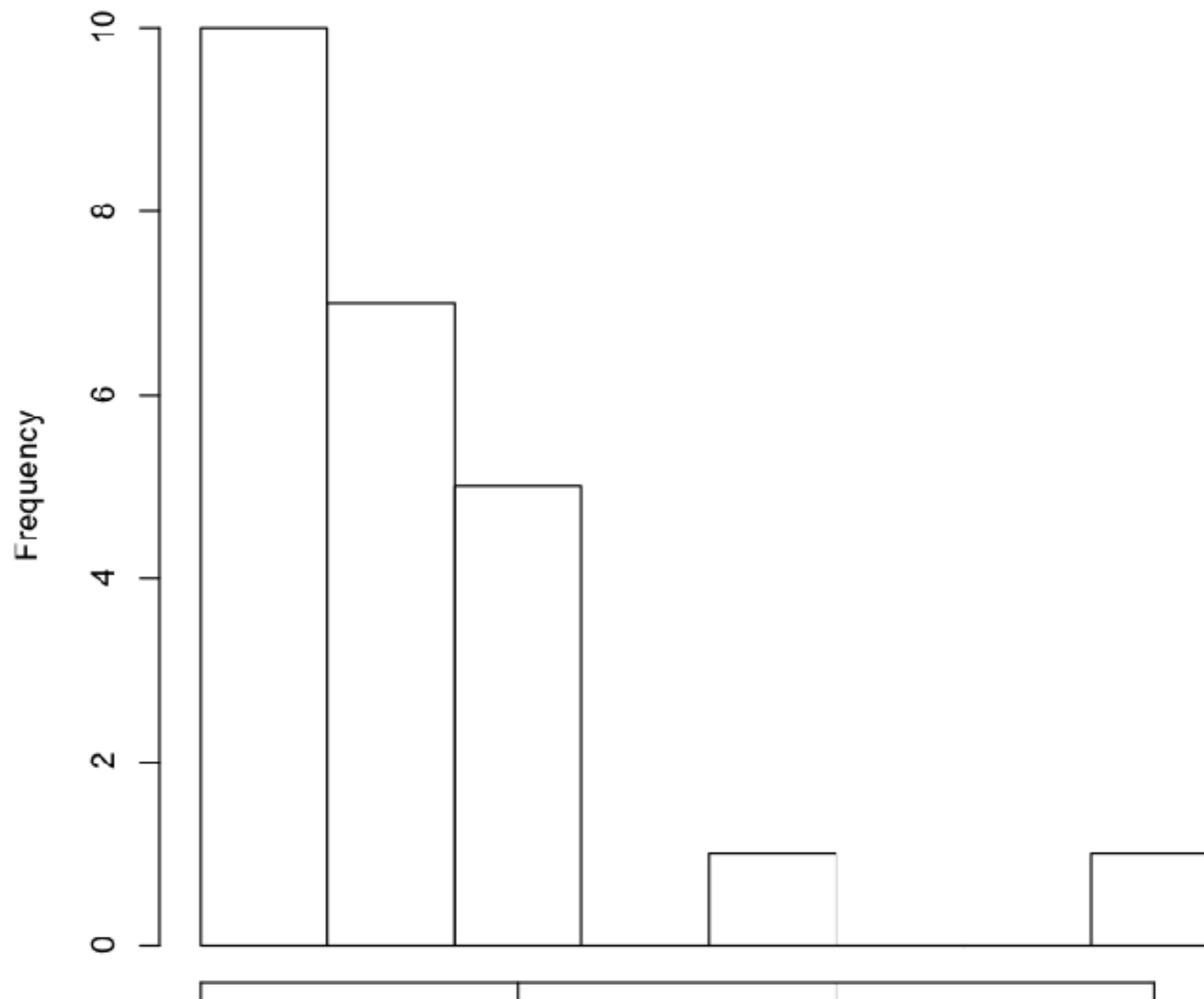


$$X_{\text{med}} = 29.000$$



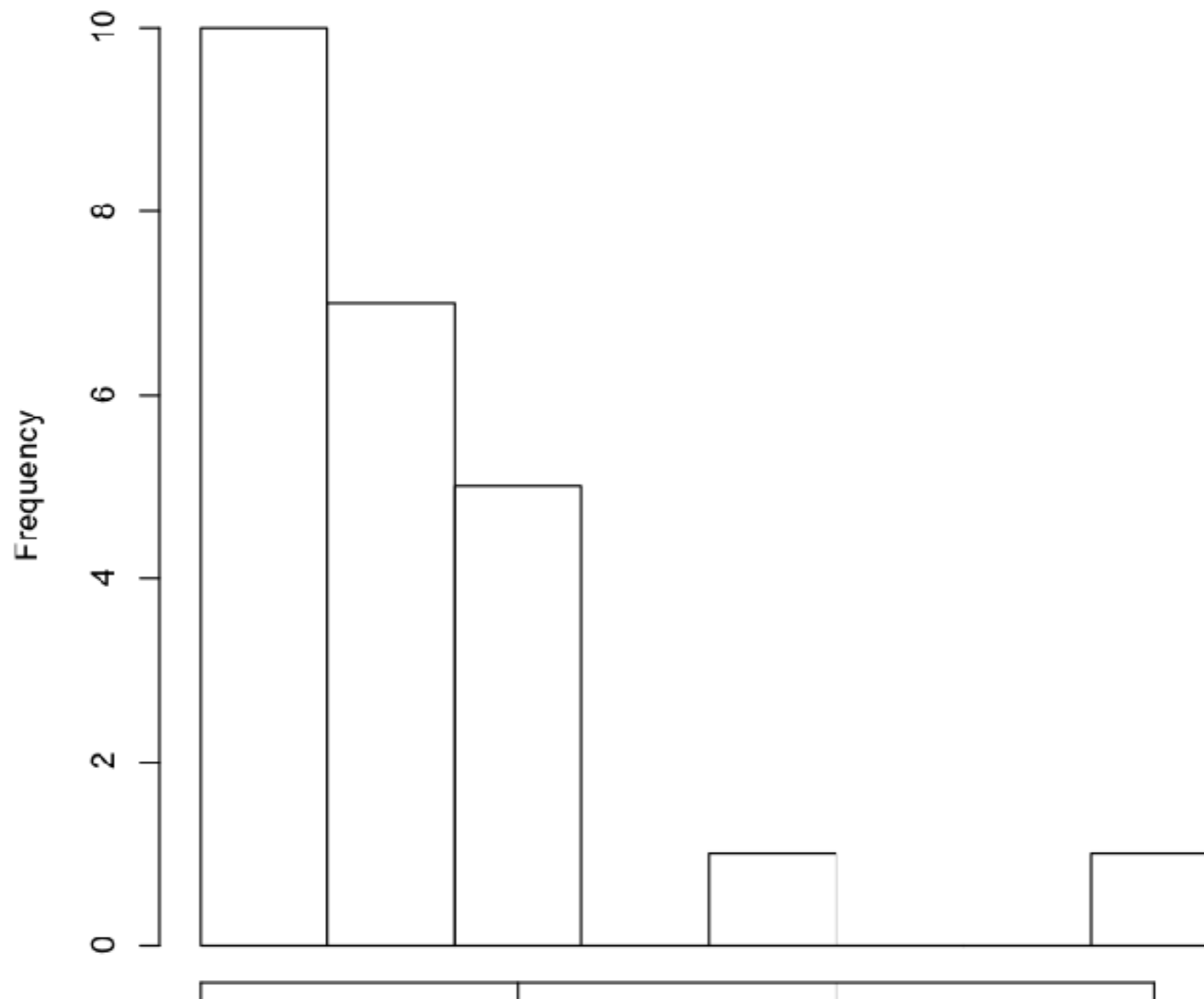
# Pohádka o Zbohatlíkově

Odstraníme na chvíli extrémní (odlehlou) hodnotu :



# Pohádka o Zbohatlíkově

Odstraníme na chvíli extrémní (odlehlou) hodnotu :



# Pohádka o Zbohatlíkově

Údaje o ročním příjmu 25 rodin ze Zbohatlíkova, n je počet členů domácnosti:

| roční příjem | n | roční příjem | n | roční příjem | n |        |   |
|--------------|---|--------------|---|--------------|---|--------|---|
| 1,200.000    | 3 | 60.000       | 1 | 45.000       | 2 | 29.000 | 3 |
| 150.000      | 5 | 51.000       | 3 | 42.000       | 2 | 26.000 | 4 |
| 86.000       | 4 | 49.000       | 4 | 38.000       | 4 | 24.000 | 4 |
| 37.000       | 3 | 20.000       | 7 | 14.000       | 1 | 18.000 | 4 |
| 35.000       | 5 | 18.000       | 3 | 13.000       | 4 | 16.000 | 3 |
| 32.000       | 3 | 18.000       | 8 | 11.000       | 1 | 16.000 | 2 |
|              |   |              |   |              |   | 10.000 | 2 |

Odstraníme na chvíli dvě extrémní (odlehlé) hodnoty :





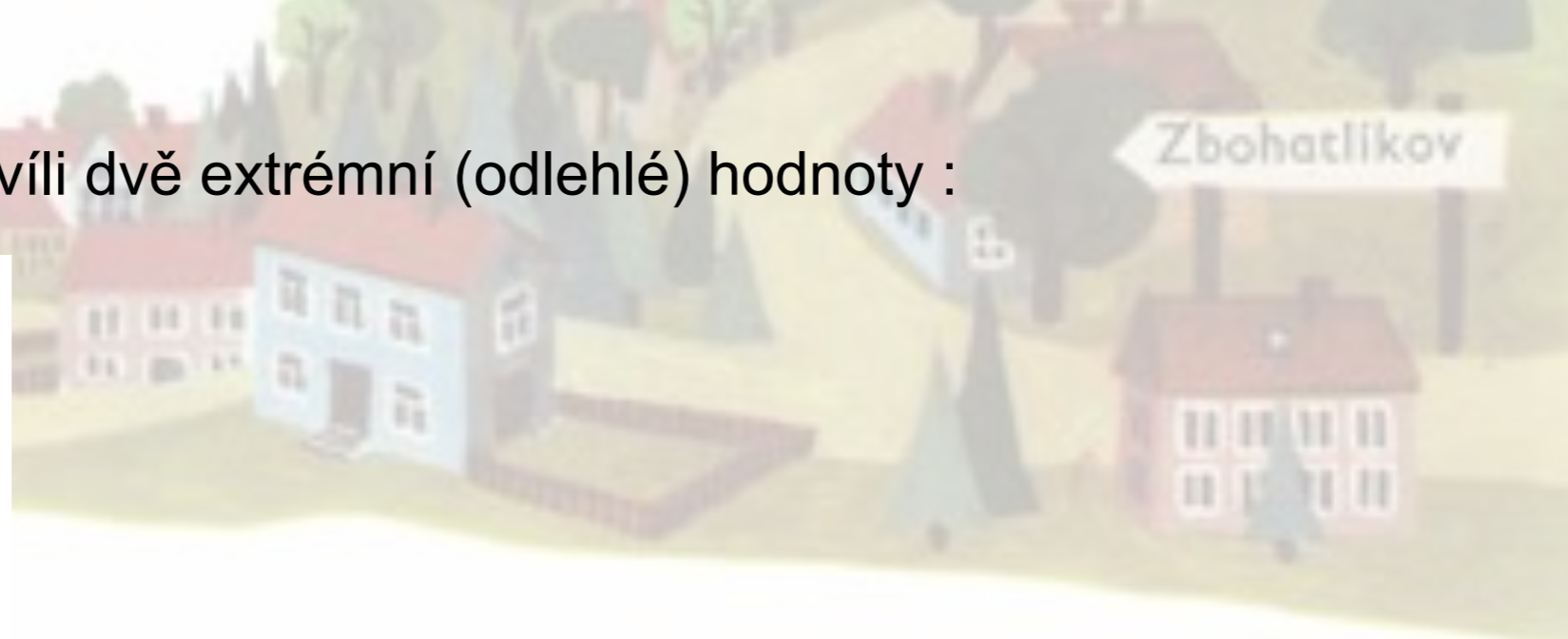
# Pohádka o Zbohatlíkově

Údaje o ročním příjmu 25 rodin ze Zbohatlíkova, n je počet členů domácnosti:

| roční příjem | n | roční příjem | n | roční příjem | n |        |   |
|--------------|---|--------------|---|--------------|---|--------|---|
| 1,200.000    | 3 | 60.000       | 1 | 45.000       | 2 | 29.000 | 3 |
| 150.000      | 5 | 51.000       | 3 | 42.000       | 2 | 26.000 | 4 |
| 86.000       | 4 | 49.000       | 4 | 38.000       | 4 | 24.000 | 4 |
| 37.000       | 3 | 20.000       | 7 | 14.000       | 1 | 18.000 | 4 |
| 35.000       | 5 | 18.000       | 3 | 13.000       | 4 | 16.000 | 3 |
| 32.000       | 3 | 18.000       | 8 | 11.000       | 1 | 16.000 | 2 |
|              |   |              |   |              |   | 10.000 | 2 |

Odstraníme na chvíli dvě extrémní (odlehle) hodnoty :

0 | 013466888  
2 | 04692578  
4 | 2591  
6 | 0  
8 | 6





# Pohádka o Zbohatlíkově

Údaje o ročním příjmu 25 rodin ze Zbohatlíkova, n je počet členů domácnosti:

| roční příjem | n | roční příjem | n | roční příjem | n |        |   |
|--------------|---|--------------|---|--------------|---|--------|---|
| 1,200.000    | 3 | 60.000       | 1 | 45.000       | 2 | 29.000 | 3 |
| 150.000      | 5 | 51.000       | 3 | 42.000       | 2 | 26.000 | 4 |
| 86.000       | 4 | 49.000       | 4 | 38.000       | 4 | 24.000 | 4 |
| 37.000       | 3 | 20.000       | 7 | 14.000       | 1 | 18.000 | 4 |
| 35.000       | 5 | 18.000       | 3 | 13.000       | 4 | 16.000 | 3 |
| 32.000       | 3 | 18.000       | 8 | 11.000       | 1 | 16.000 | 2 |
|              |   |              |   |              |   | 10.000 | 2 |

Odstraníme na chvíli dvě extrémní (odlehlé) hodnoty :

0 | 013466888  
2 | 04692578  
4 | 2591  
6 | 0  
8 | 6

Nejčastější hodnota  
je 18.000 tolarů

# Pohádka o Zbohatlíkově

Údaje o ročním příjmu 25 rodin ze Zbohatlíkova, n je počet členů domácnosti:

| roční příjem | n | roční příjem | n | roční příjem | n |        |   |
|--------------|---|--------------|---|--------------|---|--------|---|
| 1,200.000    | 3 | 60.000       | 1 | 45.000       | 2 | 29.000 | 3 |
| 150.000      | 5 | 51.000       | 3 | 42.000       | 2 | 26.000 | 4 |
| 86.000       | 4 | 49.000       | 4 | 38.000       | 4 | 24.000 | 4 |
| 37.000       | 3 | 20.000       | 7 | 14.000       | 1 | 18.000 | 4 |
| 35.000       | 5 | 18.000       | 3 | 13.000       | 4 | 16.000 | 3 |
| 32.000       | 3 | 18.000       | 8 | 11.000       | 1 | 16.000 | 2 |
|              |   |              |   |              |   | 10.000 | 2 |

Odstraníme na chvíli dvě extrémní (odlehlé) hodnoty :

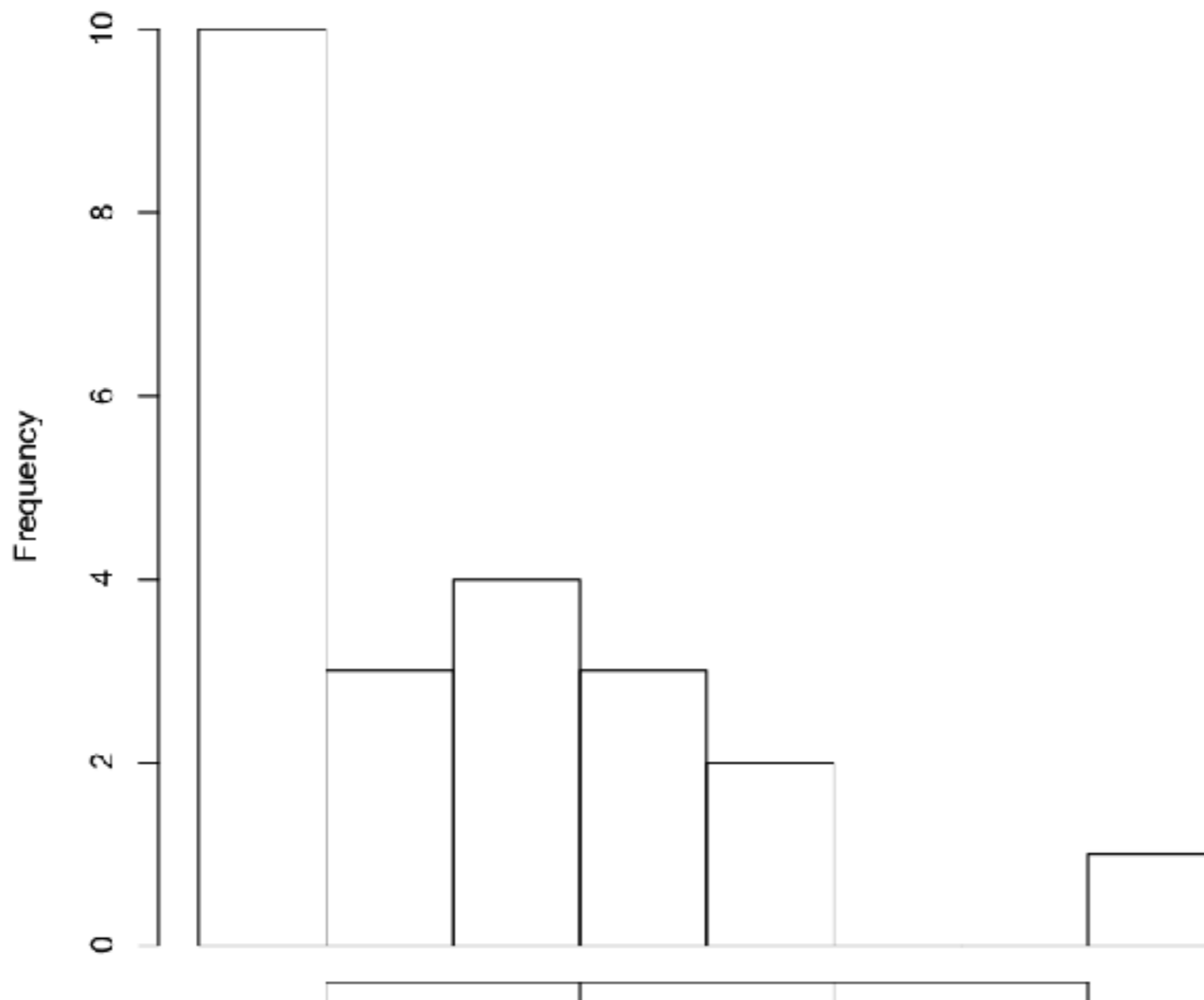
0 | 013466888  
 2 | 04692578  
 4 | 2591  
 6 | 0  
 8 | 6

Nejčastější hodnota  
 je 18.000 tolarů



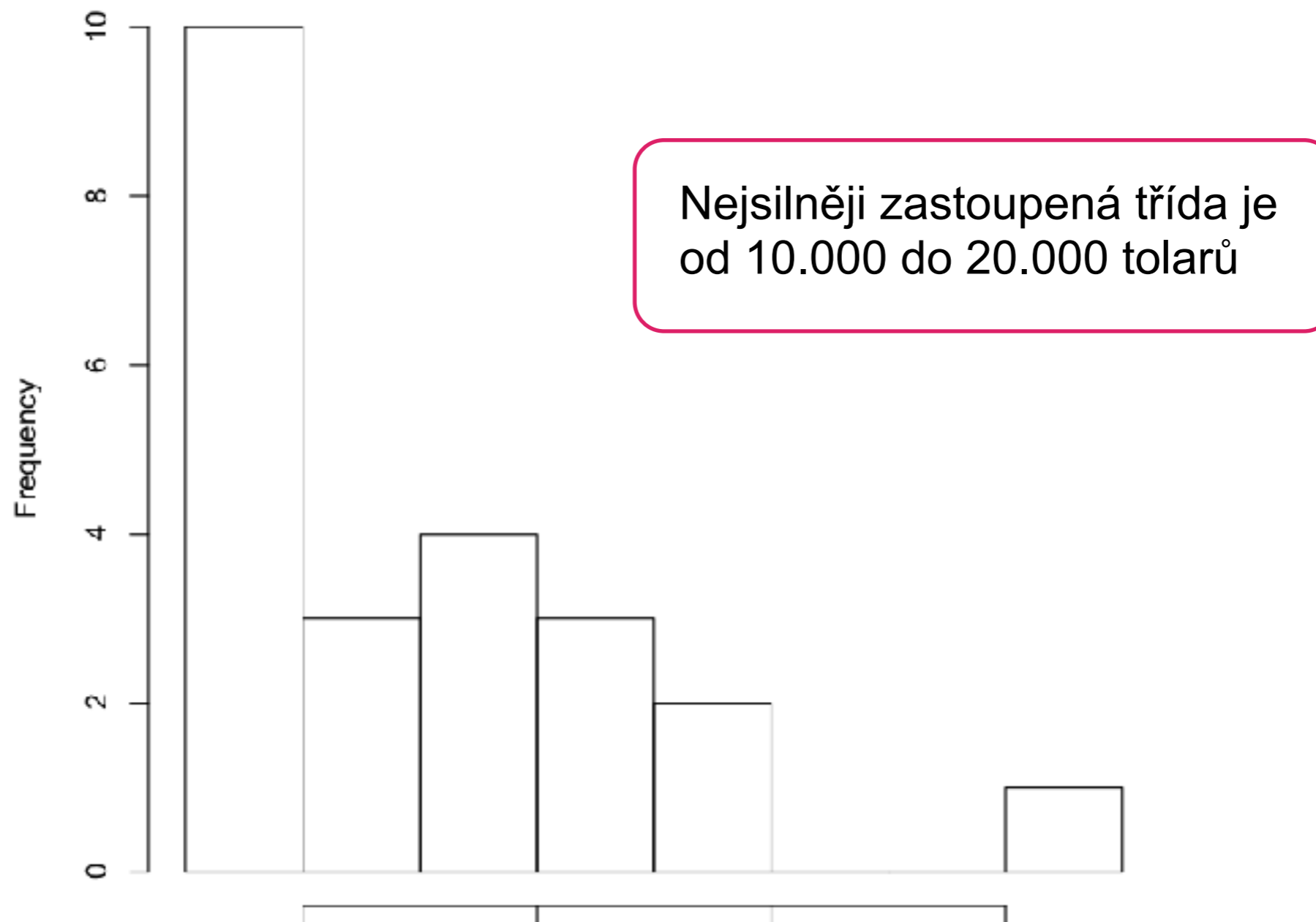
# Pohádka o Zbohatlívě

Odstraníme na chvíli dvě extrémní (odlehle) hodnoty:



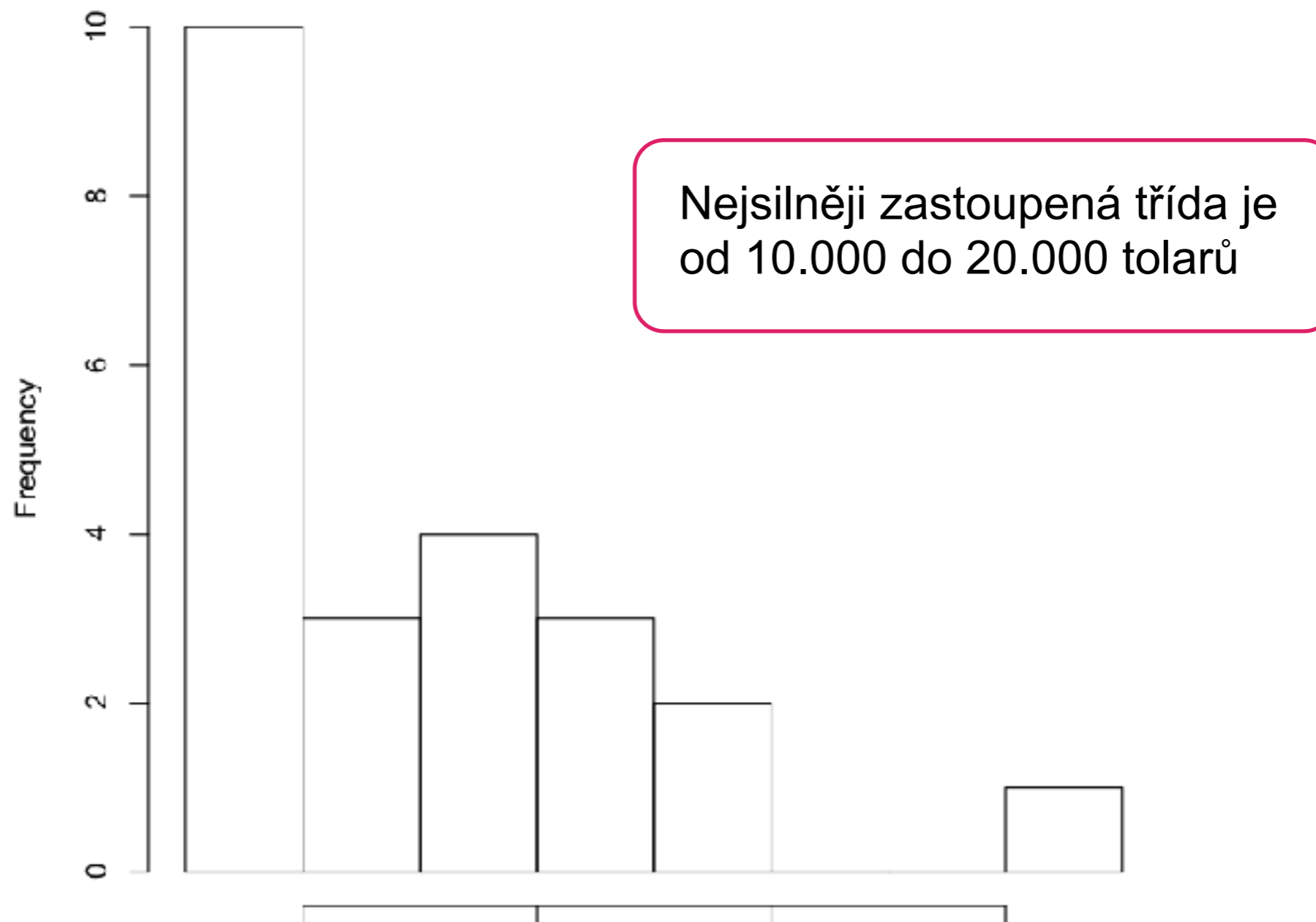
# Pohádka o Zbohatlíkově

Odstraníme na chvíli dvě extrémní (odlehle) hodnoty:



# Pohádka o Zbohatlíkově

Odstraníme na chvíli dvě extrémní (odlehle) hodnoty:



# Pohádka o Zbohatlíkově

Příjmy na hlavu (83):

|          |          |          |         |         |         |         |         |         |         |
|----------|----------|----------|---------|---------|---------|---------|---------|---------|---------|
| 400.000, | 400.000, | 400.000, | 30.000, | 30.000, | 30.000, | 30.000, | 30.000, | 21.500, | 1.500,  |
| 21.500,  | 21.500,  | 12.333,  | 12.333, | 12.333, | 7.000,  | 7.000,  | 7.000,  | 7.000,  | 7.000,  |
| 10.666,  | 10.666,  | 10.666,  | 9.666,  | 9.666,  | 9.666,  | 6.500,  | 6.500,  | 6.500,  | 6.500,  |
| 6.000,   | 6.000,   | 6.000,   | 6.000,  | 60.000, | 17.000, | 12.250, | 12.250, | 12.250, | 12.250, |
| 2.857,   | 2.857,   | 2.857,   | 2.857,  | 2.857,  | 2.857,  | 2.857,  | 6.000,  | 6.000,  | 6.000,  |
| 2.250,   | 2.250,   | 2.250,   | 2.250,  | 2.250,  | 2.250,  | 2.250,  | 2.250,  | 4.500,  | 4.500,  |
| 4.500,   | 4.500,   | 5.333,   | 5.333,  | 5.333,  | 8.000,  | 8.000,  | 22.500, | 22.500, | 21.000, |
| 21.000,  | 9.500,   | 9.500,   | 9.500,  | 9.500,  | 14.000, | 3.250,  | 3.250,  | 3.250,  | 3.250,  |
| 11.000,  | 5.000,   | 5.000    |         |         |         |         |         |         |         |



# Pohádka o Zbohatlíkově

Příjmy na hlavu (83):

|          |          |          |         |         |         |         |         |         |         |
|----------|----------|----------|---------|---------|---------|---------|---------|---------|---------|
| 400.000, | 400.000, | 400.000, | 30.000, | 30.000, | 30.000, | 30.000, | 30.000, | 21.500, | 1.500,  |
| 21.500,  | 21.500,  | 12.333,  | 12.333, | 12.333, | 7.000,  | 7.000,  | 7.000,  | 7.000,  | 7.000,  |
| 10.666,  | 10.666,  | 10.666,  | 9.666,  | 9.666,  | 9.666,  | 6.500,  | 6.500,  | 6.500,  | 6.500,  |
| 6.000,   | 6.000,   | 6.000,   | 6.000,  | 60.000, | 17.000, | 12.250, | 12.250, | 12.250, | 12.250, |
| 2.857,   | 2.857,   | 2.857,   | 2.857,  | 2.857,  | 2.857,  | 2.857,  | 6.000,  | 6.000,  | 6.000,  |
| 2.250,   | 2.250,   | 2.250,   | 2.250,  | 2.250,  | 2.250,  | 2.250,  | 2.250,  | 4.500,  | 4.500,  |
| 4.500,   | 4.500,   | 5.333,   | 5.333,  | 5.333,  | 8.000,  | 8.000,  | 22.500, | 22.500, | 21.000, |
| 21.000,  | 9.500,   | 9.500,   | 9.500,  | 9.500,  | 14.000, | 3.250,  | 3.250,  | 3.250,  | 3.250,  |
| 11.000,  | 5.000,   | 5.000    |         |         |         |         |         |         |         |

Uspořádané příjmy na hlavu:

|         |         |         |        |        |        |        |        |        |        |
|---------|---------|---------|--------|--------|--------|--------|--------|--------|--------|
| 2.250   | 2.250   | 2.250   | 2.250  | 2.250  | 2.250  | 2.250  | 2.250  | 2.857  | 2.857  |
| 2.857   | 2.857   | 2.857   | 2.857  | 2.857  | 3.250  | 3.250  | 3.250  | 3.250  | 4.500  |
| 4.500   | 4.500   | 4.500   | 5.000  | 5.000  | 5.333  | 5.333  | 5.333  | 6.000  | 6.000  |
| 6.000   | 6.000   | 6.000   | 6.000  | 6.000  | 6.500  | 6.500  | 6.500  | 6.500  | 7.000  |
| 7.000   | 7.000   | 7.000   | 7.000  | 8.000  | 8.000  | 9.500  | 9.500  | 9.500  | 9.500  |
| 9.666   | 9.666   | 9.666   | 10.666 | 10.666 | 10.666 | 11.000 | 12.250 | 12.250 | 12.250 |
| 12.250  | 12.333  | 12.333  | 12.333 | 14.000 | 17.000 | 21.000 | 21.000 | 21.500 | 21.500 |
| 21.500  | 21.500  | 22.500  | 22.500 | 30.000 | 30.000 | 30.000 | 30.000 | 30.000 | 60.000 |
| 400.000 | 400.000 | 400.000 |        |        |        |        |        |        |        |



# Pohádka o Zbohatlíkově

Příjmy na hlavu (83):

|          |          |          |         |         |         |         |         |         |         |
|----------|----------|----------|---------|---------|---------|---------|---------|---------|---------|
| 400.000, | 400.000, | 400.000, | 30.000, | 30.000, | 30.000, | 30.000, | 30.000, | 21.500, | 1.500,  |
| 21.500,  | 21.500,  | 12.333,  | 12.333, | 12.333, | 7.000,  | 7.000,  | 7.000,  | 7.000,  | 7.000,  |
| 10.666,  | 10.666,  | 10.666,  | 9.666,  | 9.666,  | 9.666,  | 6.500,  | 6.500,  | 6.500,  | 6.500,  |
| 6.000,   | 6.000,   | 6.000,   | 6.000,  | 60.000, | 17.000, | 12.250, | 12.250, | 12.250, | 12.250, |
| 2.857,   | 2.857,   | 2.857,   | 2.857,  | 2.857,  | 2.857,  | 2.857,  | 6.000,  | 6.000,  | 6.000,  |
| 2.250,   | 2.250,   | 2.250,   | 2.250,  | 2.250,  | 2.250,  | 2.250,  | 2.250,  | 4.500,  | 4.500,  |
| 4.500,   | 4.500,   | 5.333,   | 5.333,  | 5.333,  | 8.000,  | 8.000,  | 22.500, | 22.500, | 21.000, |
| 21.000,  | 9.500,   | 9.500,   | 9.500,  | 9.500,  | 14.000, | 3.250,  | 3.250,  | 3.250,  | 3.250,  |
| 11.000,  | 5.000,   | 5.000    |         |         |         |         |         |         |         |

Uspořádané příjmy na hlavu: 80% je 66,4 lidí

|         |         |         |        |        |        |        |        |        |        |
|---------|---------|---------|--------|--------|--------|--------|--------|--------|--------|
| 2.250   | 2.250   | 2.250   | 2.250  | 2.250  | 2.250  | 2.250  | 2.250  | 2.857  | 2.857  |
| 2.857   | 2.857   | 2.857   | 2.857  | 2.857  | 3.250  | 3.250  | 3.250  | 3.250  | 4.500  |
| 4.500   | 4.500   | 4.500   | 5.000  | 5.000  | 5.333  | 5.333  | 5.333  | 6.000  | 6.000  |
| 6.000   | 6.000   | 6.000   | 6.000  | 6.000  | 6.500  | 6.500  | 6.500  | 6.500  | 7.000  |
| 7.000   | 7.000   | 7.000   | 7.000  | 8.000  | 8.000  | 9.500  | 9.500  | 9.500  | 9.500  |
| 9.666   | 9.666   | 9.666   | 10.666 | 10.666 | 10.666 | 11.000 | 12.250 | 12.250 | 12.250 |
| 12.250  | 12.333  | 12.333  | 12.333 | 14.000 | 17.000 | 21.000 | 21.000 | 21.500 | 21.500 |
| 21.500  | 21.500  | 22.500  | 22.500 | 30.000 | 30.000 | 30.000 | 30.000 | 30.000 | 60.000 |
| 400.000 | 400.000 | 400.000 |        |        |        |        |        |        |        |

# Pohádka o Zbohatlíkově

Příjmy na hlavu (83):

400.000, 400.000, 400.000, 30.000, 30.000, 30.000, 30.000, 30.000, 21.500, 1.500,  
 21.500, 21.500, 12.333, 12.333, 12.333, 7.000, 7.000, 7.000, 7.000, 7.000,  
 10.666, 10.666, 10.666, 9.666, 9.666, 9.666, 6.500, 6.500, 6.500, 6.500,  
 6.000, 6.000, 6.000, 6.000, 60.000, 17.000, 12.250, 12.250, 12.250, 12.250,  
 2.857, 2.857, 2.857, 2.857, 2.857, 2.857, 2.857, 6.000, 6.000, 6.000,  
 2.250, 2.250, 2.250, 2.250, 2.250, 2.250, 2.250, 2.250, 4.500, 4.500,  
 4.500, 4.500, 5.333, 5.333, 5.333, 8.000, 8.000, 22.500, 22.500, 21.000,  
 21.000, 9.500, 9.500, 9.500, 9.500, 14.000, 3.250, 3.250, 3.250, 3.250,  
 11.000, 5.000, 5.000

Uspořádané příjmy na hlavu: 80% je 66,4 lidí

|         |         |         |        |        |        |        |        |        |        |
|---------|---------|---------|--------|--------|--------|--------|--------|--------|--------|
| 2.250   | 2.250   | 2.250   | 2.250  | 2.250  | 2.250  | 2.250  | 2.250  | 2.857  | 2.857  |
| 2.857   | 2.857   | 2.857   | 2.857  | 2.857  | 3.250  | 3.250  | 3.250  | 3.250  | 4.500  |
| 4.500   | 4.500   | 4.500   | 5.000  | 5.000  | 5.333  | 5.333  | 5.333  | 6.000  | 6.000  |
| 6.000   | 6.000   | 6.000   | 6.000  | 6.000  | 6.500  | 6.500  | 6.500  | 6.500  | 7.000  |
| 7.000   | 7.000   | 7.000   | 7.000  | 8.000  | 8.000  | 9.500  | 9.500  | 9.500  | 9.500  |
| 9.666   | 9.666   | 9.666   | 10.666 | 10.666 | 10.666 | 11.000 | 12.250 | 12.250 | 12.250 |
| 12.250  | 12.333  | 12.333  | 12.333 | 14.000 | 17.000 | 21.000 | 21.000 | 21.500 | 21.500 |
| 21.500  | 21.500  | 22.500  | 22.500 | 30.000 | 30.000 | 30.000 | 30.000 | 30.000 | 60.000 |
| 400.000 | 400.000 | 400.000 |        |        |        |        |        |        |        |

# Pohádka o Zbohatlíkově

Příjmy na hlavu (83):

400.000, 400.000, 400.000, 30.000, 30.000, 30.000, 30.000, 30.000, 21.500, 1.500,  
 21.500, 21.500, 12.333, 12.333, 12.333, 7.000, 7.000, 7.000, 7.000, 7.000,  
 10.666, 10.666, 10.666, 9.666, 9.666, 9.666, 6.500, 6.500, 6.500, 6.500,  
 6.000, 6.000, 6.000, 6.000, 60.000, 17.000, 12.250, 12.250, 12.250, 12.250,  
 2.857, 2.857, 2.857, 2.857, 2.857, 2.857, 2.857, 6.000, 6.000, 6.000,  
 2.250, 2.250, 2.250, 2.250, 2.250, 2.250, 2.250, 2.250, 4.500, 4.500,  
 4.500, 4.500, 5.333, 5.333, 5.333, 8.000, 8.000, 22.500, 22.500, 21.000,  
 21.000, 9.500, 9.500, 9.500, 9.500, 14.000, 3.250, 3.250, 3.250, 3.250,  
 11.000, 5.000, 5.000

Uspořádané příjmy na hlavu: 80% je 66,4 lidí

|         |         |         |        |        |        |        |        |        |        |
|---------|---------|---------|--------|--------|--------|--------|--------|--------|--------|
| 2.250   | 2.250   | 2.250   | 2.250  | 2.250  | 2.250  | 2.250  | 2.250  | 2.857  | 2.857  |
| 2.857   | 2.857   | 2.857   | 2.857  | 2.857  | 3.250  | 3.250  | 3.250  | 3.250  | 4.500  |
| 4.500   | 4.500   | 4.500   | 5.000  | 5.000  | 5.333  | 5.333  | 5.333  | 6.000  | 6.000  |
| 6.000   | 6.000   | 6.000   | 6.000  | 6.000  | 6.500  | 6.500  | 6.500  | 6.500  | 7.000  |
| 7.000   | 7.000   | 7.000   | 7.000  | 8.000  | 8.000  | 9.500  | 9.500  | 9.500  | 9.500  |
| 9.666   | 9.666   | 9.666   | 10.666 | 10.666 | 10.666 | 11.000 | 12.250 | 12.250 | 12.250 |
| 12.250  | 12.333  | 12.333  | 12.333 | 14.000 | 17.000 | 21.000 | 21.000 | 21.500 | 21.500 |
| 21.500  | 21.500  | 22.500  | 22.500 | 30.000 | 30.000 | 30.000 | 30.000 | 30.000 | 60.000 |
| 400.000 | 400.000 | 400.000 |        |        |        |        |        |        |        |

**80% lidí má menší roční příjem než 20.000 tolarů**

# Pohádka o Zbohatlíkově

Příjmy na hlavu (83):

400.000, 400.000, 400.000, 30.000, 30.000, 30.000, 30.000, 30.000, 21.500, 1.500,  
 21.500, 21.500, 12.333, 12.333, 12.333, 7.000, 7.000, 7.000, 7.000, 7.000,  
 10.666, 10.666, 10.666, 9.666, 9.666, 9.666, 6.500, 6.500, 6.500, 6.500,  
 6.000, 6.000, 6.000, 6.000, 60.000, 17.000, 12.250, 12.250, 12.250, 12.250,  
 2.857, 2.857, 2.857, 2.857, 2.857, 2.857, 2.857, 6.000, 6.000, 6.000,  
 2.250, 2.250, 2.250, 2.250, 2.250, 2.250, 2.250, 2.250, 4.500, 4.500,  
 4.500, 4.500, 5.333, 5.333, 5.333, 8.000, 8.000, 22.500, 22.500, 21.000,  
 21.000, 9.500, 9.500, 9.500, 9.500, 14.000, 3.250, 3.250, 3.250, 3.250,  
 11.000, 5.000, 5.000

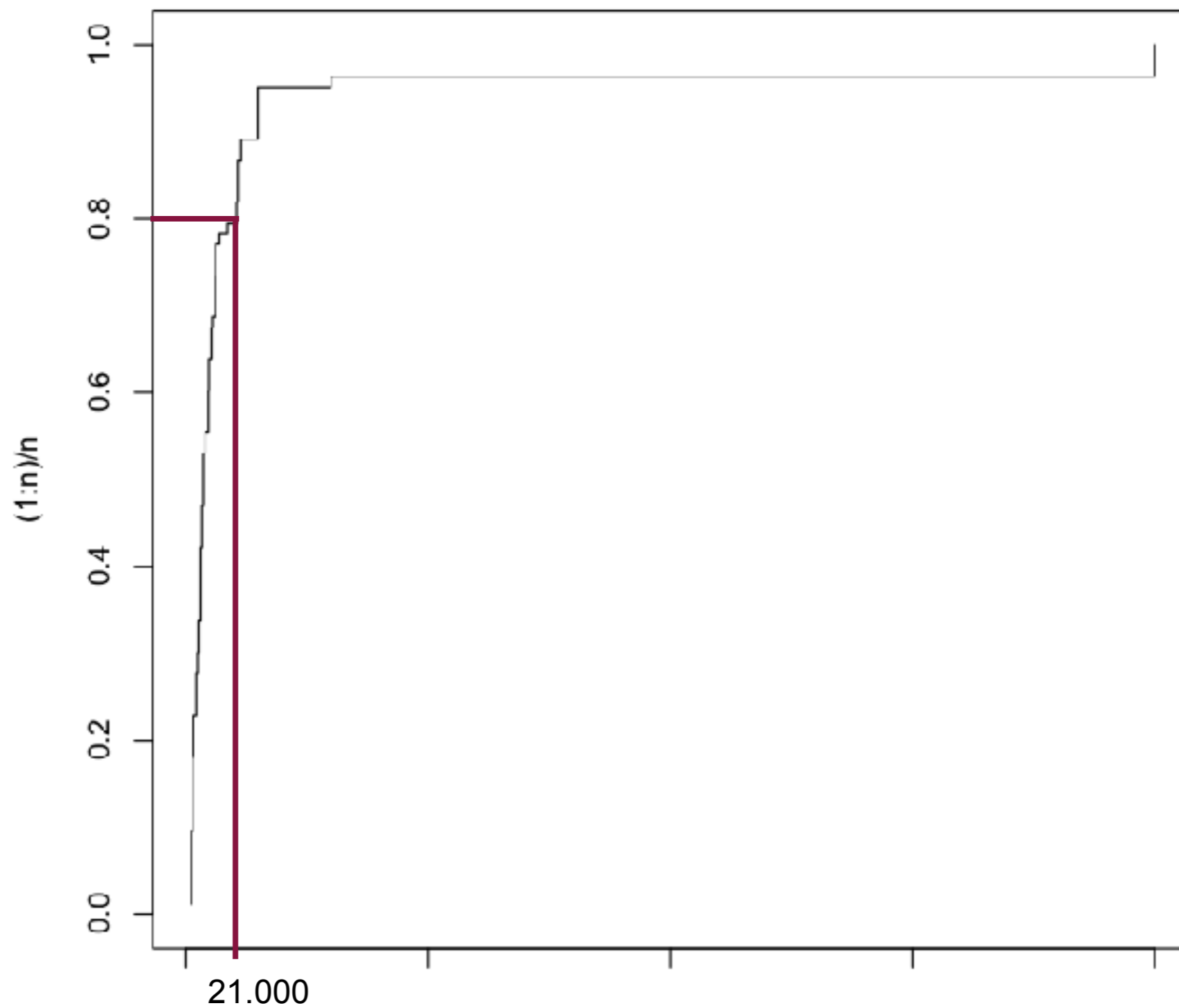
Uspořádané příjmy na hlavu: 80% je 66,4 lidí

|         |         |         |        |        |        |        |        |        |        |
|---------|---------|---------|--------|--------|--------|--------|--------|--------|--------|
| 2.250   | 2.250   | 2.250   | 2.250  | 2.250  | 2.250  | 2.250  | 2.250  | 2.857  | 2.857  |
| 2.857   | 2.857   | 2.857   | 2.857  | 2.857  | 3.250  | 3.250  | 3.250  | 3.250  | 4.500  |
| 4.500   | 4.500   | 4.500   | 5.000  | 5.000  | 5.333  | 5.333  | 5.333  | 6.000  | 6.000  |
| 6.000   | 6.000   | 6.000   | 6.000  | 6.000  | 6.500  | 6.500  | 6.500  | 6.500  | 7.000  |
| 7.000   | 7.000   | 7.000   | 7.000  | 8.000  | 8.000  | 9.500  | 9.500  | 9.500  | 9.500  |
| 9.666   | 9.666   | 9.666   | 10.666 | 10.666 | 10.666 | 11.000 | 12.250 | 12.250 | 12.250 |
| 12.250  | 12.333  | 12.333  | 12.333 | 14.000 | 17.000 | 21.000 | 21.000 | 21.500 | 21.500 |
| 21.500  | 21.500  | 22.500  | 22.500 | 30.000 | 30.000 | 30.000 | 30.000 | 30.000 | 60.000 |
| 400.000 | 400.000 | 400.000 |        |        |        |        |        |        |        |

**80% lidí má menší roční příjem než 20.000 tolarů** →

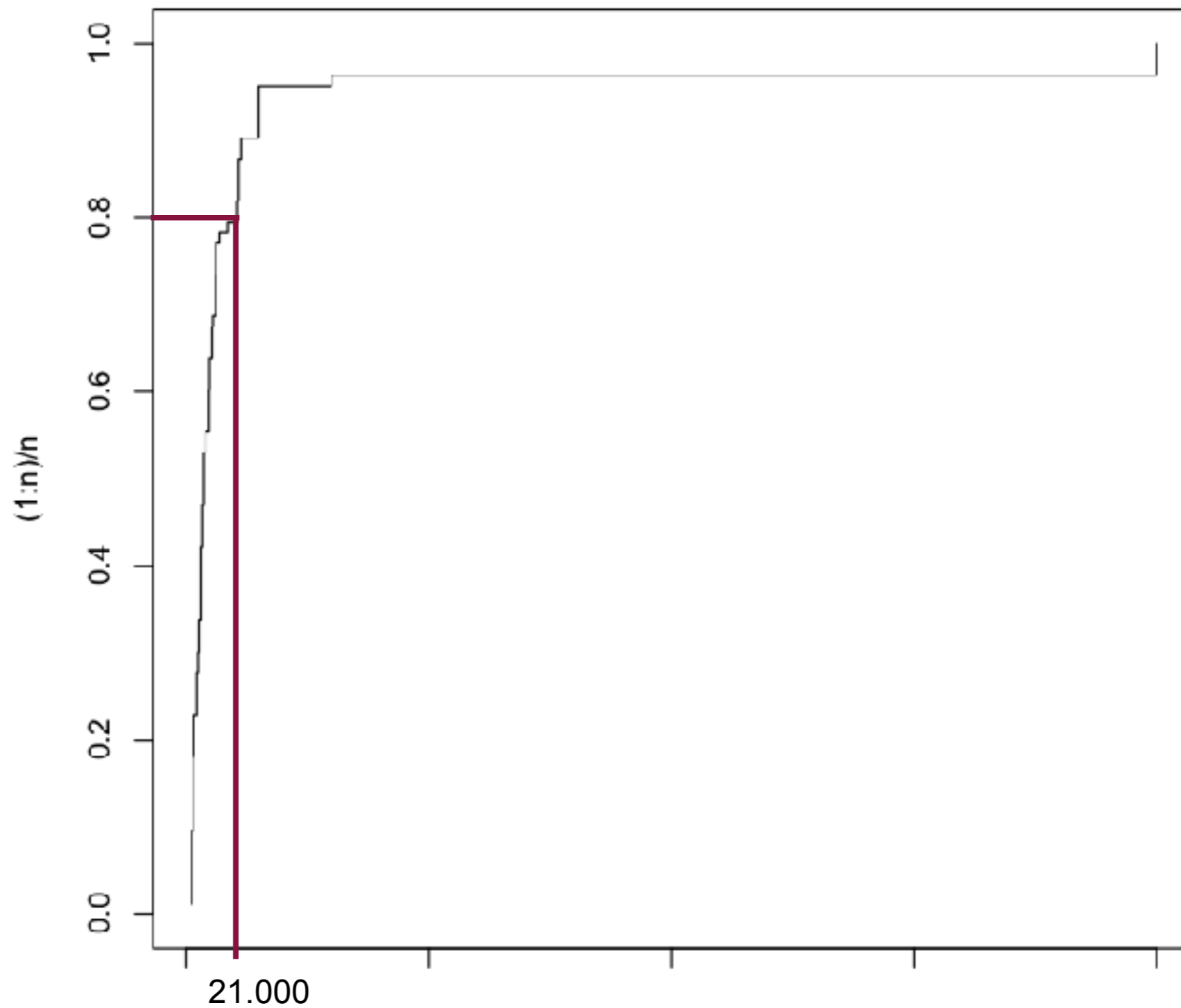
# Pohádka o Zbohatlívě

Empirická distribuční funkce:



# Pohádka o Zbohatlíkově

Empirická distribuční funkce:





# Pohádka o Zbohatlíkově

Lhal tedy zprostředkovatel?





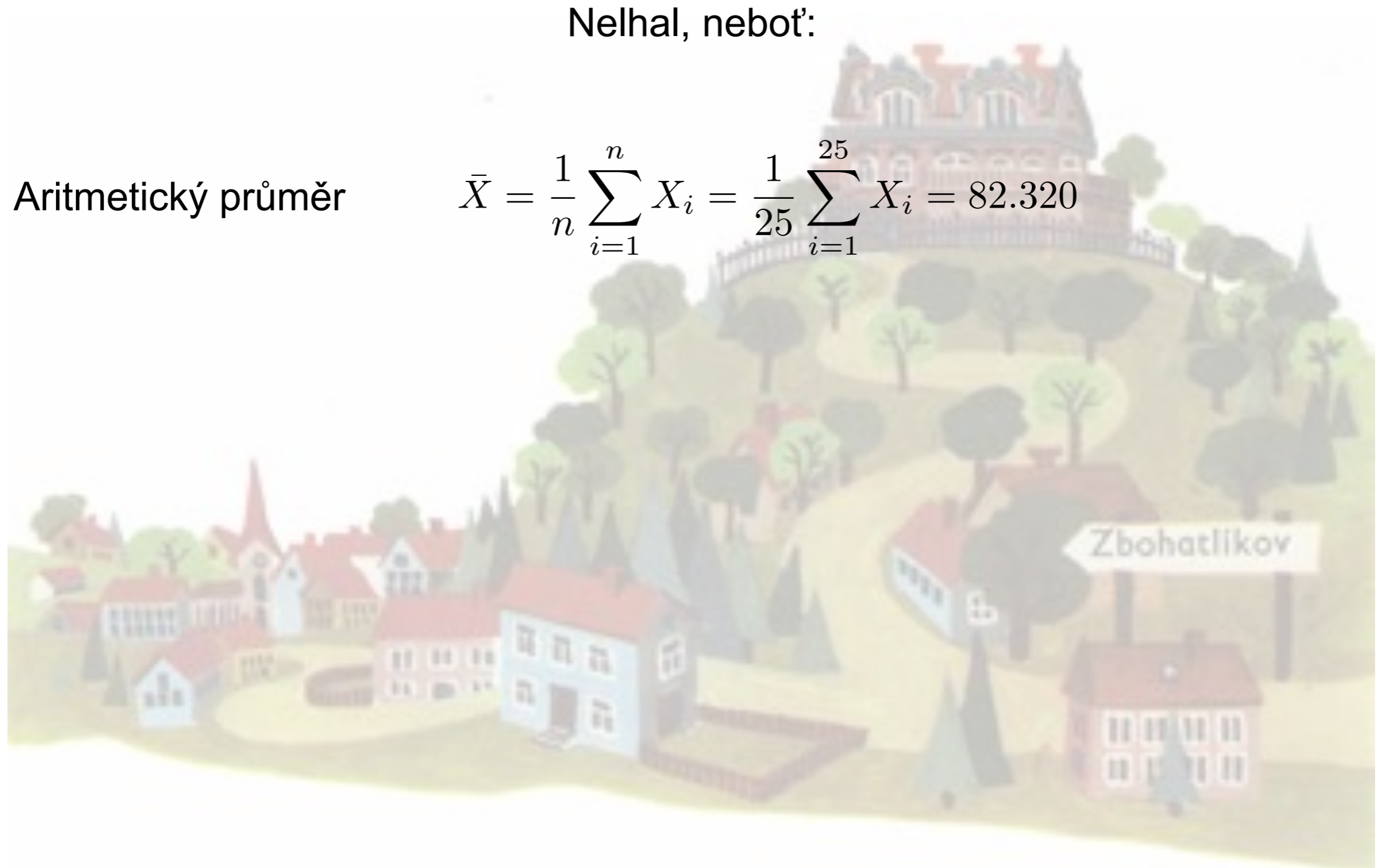
# Pohádka o Zbohatlíkově

Lhal tedy zprostředkovatel?

Nelhal, neboť:

Aritmetický průměr

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{25} \sum_{i=1}^{25} X_i = 82.320$$



# Pohádka o Zbohatlíkově

Lhal tedy zprostředkovatel?

Nelhal, neboť:

Aritmetický průměr

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{25} \sum_{i=1}^{25} X_i = 82.320$$

V případě příjmů je však lépe použít:

Geometrický průměr

$$\hat{X} = \left( \prod_{i=1}^n X_i \right)^{\frac{1}{n}} = 32.730$$

(Neboť mzdy mají zpravidla silně sešikmené rozdělení)



# Pohádka o Zbohatlíkově

Lhal tedy zprostředkovatel?

Nelhal, neboť:

Aritmetický průměr

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{25} \sum_{i=1}^{25} X_i = 82.320$$

V případě příjmů je však lépe použít:

Geometrický průměr

$$\hat{X} = \left( \prod_{i=1}^n X_i \right)^{\frac{1}{n}} = 32.730$$

(Neboť mzdy mají zpravidla silně sešikmené rozdělení)

