

# Popisná statistika

## Základní pojmy

Naměřené hodnoty:

$$x_1, x_2, \dots, x_n,$$

Seřazené naměřené hodnoty:

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

Aritmetický průměr

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

**p-quantil** (výběrový p-quantil)

$$\tilde{x}_p = \begin{cases} x_{([np]+1)} & np \neq [np] \\ \frac{1}{2}(x_{(np)} + x_{(np+1)}) & np = [np] \end{cases},$$

kde  $[a]$  značí celou část z  $a$  a  $0 < p < 1$ .

Speciální případy:

- 0,5-quantil = **Medián** – 50 % hodnot leží „pod“ touto hodnotou a 50 % „nad“
- 0,25-quantil resp. 0,75-quantil = **dolní** resp. **horní kvartil**
- 0,1-quantil, 0,2-quantil, ... = **decily**

**Rozptyl**

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left( \sum_{i=1}^n x_i^2 - n\bar{x}^2 \right)$$

**Směrodatná odchylka** – odmocnina z rozptylu  $s = \sqrt{s^2}$

**Modus** – nejčastější hodnota, značí se  $\hat{x}$ .

**Mezikvartilové rozpětí**

$$Q = \tilde{x}_{0,75} - \tilde{x}_{0,25}$$

**Krabicový diagram** (box plot, box and whisker plot, vousatá krabice) – „Krabice“ je ohraničena hodnotami kvartilů a je zobrazen medián. „Vousky“ znázorňují hodnoty, které nejsou od jednotlivých kvartilů vzdálené o více jak 1,5 násobek  $Q$ . Jednotlivě jsou vyznačena pozorování, která jsou ve větší vzdálenosti.

1. DÚ: Byly naměřeny hodnoty nějakého jevu:

$$10; 7; 7; 8; 8; 9; 10; 9; 4; 9; 10; 9; 11; 9; 7; 8; 3; 9; 8; 7$$

Určete průměr, medián, kvartily, rozptyl, mezikvartilové rozpětí, hodnoty názorně pomocí krabicového diagramu.

$$[\bar{x} = 8,1, \hat{x} = 9, \tilde{x}_{0,5} = 8,5, \tilde{x}_{0,25} = 7, \tilde{x}_{0,75} = 9]$$

