

# Statistika

## *Principy statistického šetření:*



Statistický soubor:

Statistická jednotka:

Rozsah souboru:

Sledovaný statistický znak:

Stat. jednotka	Hodnota statistického znaku $[x_i]$
1.	
2.	
3.	
4.	
5.	
·	·
·	·
·	·
·	·
·	·
·	·
·	·
·	·
·	·
·	·
·	·
·	·

### 1. Statistický znak:

Hodnota znaku $[x_i]$	Absolutní četnost $[n_i]$	Relativní četnost $[v_i]$ (v %)

Aritmetický průměr:  $\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$

Modus:  $mod(x)$

Medián:  $med(x)$

Variační rozpětí:  $R = x_{max} - x_{min}$

Průměrná absolutní odchylka:  $\bar{d} = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$

Názvosloví:

**Statistický soubor** – množina – skupina prvků (objektů, osob, ...), které mají společné vlastnosti, např. všichni žáci třídy, zaměstnanci jedné firmy apod.

**Statistická jednotka** – prvek statistického souboru (žák, zaměstnanec, ...)

**Rozsah souboru** – počet jednotek statistického souboru (označení **n**)

Každá statistická jednotka je nositelem nějakých vlastností – **statistický znak** (vlastnosti důležité pro statistické zkoumání), např. číselné údaje o žácích třídy – výška, váha, hod kriketovým míčkem, délka chodidla, počet sourozenců, výsledky písemných prací apod.

**Hodnotu statistického znaku** pro danou statistickou jednotku značíme  $x_i$

**Absolutní četnost hodnoty znaku** (označení  $n_i$ ) – číslo, které udává, kolikrát se daná hodnota znaku ve statistickém souboru vyskytuje. Platí, že

$$n_1 + n_2 + \dots + n_k = n,$$

tzn. součet jednotlivých absolutních četností sledovaného znaku je roven rozsahu souboru.

**Relativní četnost hodnoty znaku** (označení  $v_i$ ) – poměr absolutní četnosti a rozsahu souboru pro danou hodnotu znaku. Součet relativních četností sledovaného znaku je roven 1, tj.

$$v_1 + v_2 + \dots + v_k = 1.$$

Relativní četnost hodnoty znaku je možno též uvést v %.

**Aritmetický průměr** – podíl součtu hodnot znaku zjištěných u všech jednotek souboru a počtu všech jednotek.

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

**Modus** znaku  $x$  je hodnota s největší četností, značíme  $\text{Mod}(x)$ . Udává, který výsledek je zastoupen nejvíce.

**Medián** – prostřední hodnota znaku  $x$ , jsou-li hodnoty uspořádány podle velikosti. Značíme  $\text{Med}(x)$

**Variační rozpětí** – dáno rozdílem mezi největší a nejmenší hodnotou znaku ve statistickém souboru.

$$R = x_{max} - x_{min}$$

**Absolutní odchylka** – absolutní hodnota rozdílu hodnoty znaku  $x_i$  a aritmetického průměru  $\bar{x}$ .

$$|x_i - \bar{x}|$$

**Průměrná absolutní odchylka** – podíl součtu absolutních odchylek u všech jednotek a rozsahu souboru

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

