

# *Základné pojmy štatistiky*

Teória

# ZÁKLADNÉ POJMY

- **štatistický súbor (základný súbor)** - množina, skupina prvkov, ktoré sú predmetom štatistického skúmania a pozorovania, a ktoré majú rovnaké vlastnosti z hľadiska pozorovania (napr. obyvatelia Slovenska, naša trieda,...)
- **štatistická jednotka** - prvok štatistického súboru - objekt pozorovania
- **rozsah štatistického súboru** - počet prvkov štatistického súboru, označuje sa  $N$
- **štatistický znak** - určitá vlastnosť prvkov štatistického súboru, ktorú sledujeme pri štatistickej práci
  - **kvantitatívny** - charakterizujú vlastnosti vyjadrené číselne (napr. vek, výška, hmotnosť,...)
  - **kvalitatívny** - charakterizujú vlastnosti opísané slovom resp. definíciou (napr. farba, obľúbený kvet, pohlavie,...)
- **výberový súbor** - podmnožina základného súboru, do ktorej sa vyberajú štatistické jednotky podľa vopred určených zásad. Medzi ne patrí **náhodnosť** (každý má rovnakú šancu byť vybraný) a **nezávislosť** (výber jedného neovplyvní výber druhého).
- **rozsah výberového súboru** - vždy podstatne menšie než  $N$ , označuje sa  $n$  ( $n \ll N$ )

# ZÁKLADNÉ POJMY

- hodnota štatistického znaku ( $x_i$ ) - vyjadruje poradie pozorovania
- Početnosť znaku
  - absolútna početnosť ( $n_i$ ) - udáva, koľkokrát sa v súbore vyskytuje hodnota znaku. Súčet absolútnych početností je  $n$ .
  - relatívna početnosť ( $\frac{n_i}{n}$ ) - podiel absolútnej početnosti znaku a rozsahu výberového súboru (desatinné číslo, resp.  $\times 100$  - percento). Súčet relatívnych početností je 1, resp. 100%).
- tabuľka početností (rozdelenie početností) - zaznamenanie prehľadne zistených údajov do tabuľky

# CHARAKTERISTIKY POLOHY

- **aritmetický priemer** - súčet jednotlivých hodnôt delíme ich počtom

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \dots + x_n}{n}$$

- **vážený aritmetický priemer** - namerané hodnoty sa viackrát opakujú

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k x_i \cdot n_i = \frac{x_1 \cdot n_1 + x_2 \cdot n_2 + \dots + x_k \cdot n_k}{n}$$

- **modus (najpočetnejšia hodnota) ( $\hat{x}$ , alebo  $mod(x)$ )** - výsledok, ktorý je v nameraných hodnotách zastúpený najčastejšie
- **medián (prostredná hodnota, centrálna hodnota) ( $\tilde{x}$ , alebo  $med(x)$ )**
  - prostredná hodnota usporiadaných nameraných hodnôt podľa veľkosti
  - počet pozorovaní nepárny -  $\tilde{x} = x_i$ , kde  $i = \frac{n+1}{2}$
  - počet pozorovaní párny -  $\tilde{x} = \frac{x_i + x_{i+1}}{2}$ , kde  $i = \frac{n}{2}$