

Štatistika

Štatistika je odbor, ktorý sa zaoberá zhromažďovaním, analýzou a interpretáciou údajov získaných z pozorovaní alebo experimentov.

- **štatistický súbor** je konečná neprázdna množina M ; počet všetkých prvkov množiny M nám určuje rozsah štatistického súboru a označujeme ho n
- ľubovoľná funkcia x , ktorá množine M priradí hodnoty z množiny reálnych čísel $x: M \rightarrow R$ nazývame **štatistický znak**; hodnoty štatistického znaku $x_1, x_2, x_3, \dots, x_n$; napr. výška, veľkosť, počet voličov jednotlivých strán, ...
*existuje aj štatistika na základe kvalitatívnych znakov, ale tá sa vykonáva inak

- početnosť hodnôt znaku:
 - **absolútna početnosť** hodnôt znaku x je číslo n_i , ktoré udáva, koľkokrát sa vyskytuje hodnota x_i v štatistickom súbore
 - **relatívna početnosť** hodnôt znaku x_i je číslo, ktoré udáva pomerné zastúpenie jednotlivých hodnôt v súbore, teda $\frac{n_i}{n}$ - percentuálne zastúpenie
- v štatistickom súbore definujeme:
 - aritmetický priemer znaku x :

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

*vážený aritmetický priemer:

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

- geometrický priemer znaku x :

$$\bar{x}_G = \sqrt[n]{x_1 x_2 \dots x_n}$$

*vážený geometrický priemer:

$$\bar{x}_G = \sqrt[n]{x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}}$$

- harmonický priemer znaku x :

$$\bar{x}_H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}$$

*vážený harmonický priemer:

$$\bar{x}_H = \frac{n}{\frac{n_1}{x_1} + \frac{n_2}{x_2} + \dots + \frac{n_k}{x_k}}$$

- **modus** štatistického súboru je najčastejšie sa vyskytujúca hodnota štatistického súboru, teda hodnota s najväčšou početnosťou; ozn. $Mod(x)$
- **medián** štatistického súboru je prostredná z hodnôt $x_1, x_2, x_3, \dots, x_n$, ak ich usporiadame podľa veľkosti
 - pri štatistickom súbore s nepárnym rozsahom je to prostredná hodnota, pri súbore s párnym rozsahom je to je to aritmetický priemer prostredných dvoch hodnôt $\frac{x_p + x_{p+1}}{2}$, kde $p = \frac{n}{2}$
- **smerodajná odchýlka** hodnôt štatistického súboru je číslo s , ktoré udáva, ako ďaleko sú hodnoty x_i rozmiestnené okolo aritmetického priemeru – vypočítame ju:

$$s = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

- **rozptyl (disperzia)** je druhá mocnina smerodajnej odchýlky s^2

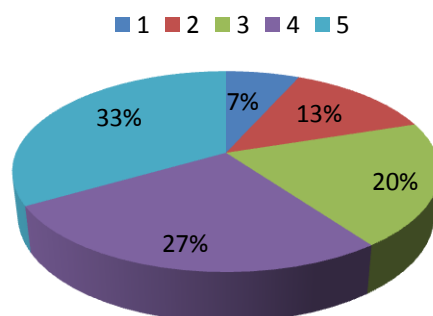
Grafické znázornenie štatistického súboru

- na príklade dosiahnutých výsledkov previerky:

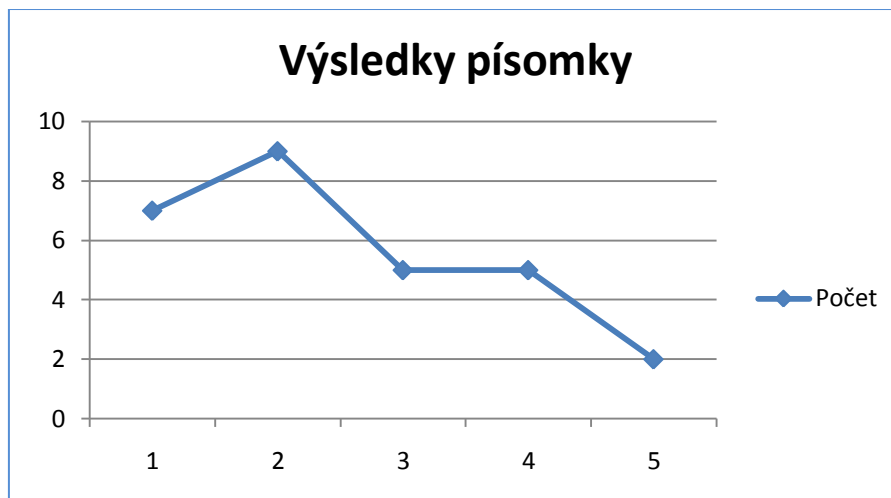
Známka	1	2	3	4	5
Počet	7	9	5	5	2

- najčastejšie používané diagramy:
- **koláčový (kruhový) diagram** – znázorňuje relatívne početnosti hodnôt znaku

Výsledky písomky



- **polygón (spojnicový diagram)** – môže požívať absolútne alebo relatívne početnosti



- **histogram (stĺpcový diagram)** - môže požívať absolútne alebo relatívne početnosti

