

4. Výběrové charakteristiky

- Míry polohy – hodnota, kolem které se data soustřeďují („střed“)
- Průměr
- Modus
- Medián
- Geometrický průměr

Průměr

- Aritmetický průměr - \bar{x}
- Kvantitativní znaky
- Nevhodný pro ordinální znaky
- Citlivý na odlehlé hodnoty
- Součet pozorování $\sum_{i=1}^n x_i = x_1 + x_2 + x_3 + \dots + x_n$
- Počet pozorování n
- Průměr $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

Měření: 39, 42, 73, 67, 24, 55

Vážený průměr

- Roztříděná data
- Vážený průměr z hodnot středů třídy (x_i)
- Vážený průměr
$$\bar{x} = \frac{1}{n} \sum_{i=1}^k n_i x_i$$

| Střed třídy x_i | Absolutní četnost n_i | Kumulativní absolutní četnost | Relativní četnost n_i/n | Kumulativní relativní četnost |
|----------------------|-------------------------|----------------------------------|---------------------------|----------------------------------|
| 120 | 13 | 13 | 0,0040 | 0,0040 |
| 125 | 95 | 108 | 0,0294 | 0,0334 |
| 130 | 414 | 522 | 0,1281 | 0,1615 |
| 135 | 880 | 1402 | 0,2724 | 0,4339 |
| 140 | 1013 | 2415 | 0,3135 | 0,7474 |
| 145 | 582 | 2997 | 0,1801 | 0,9275 |
| 150 | 199 | 3196 | 0,0616 | 0,9891 |
| 155 | 29 | 3225 | 0,0090 | 0,9981 |
| 160 | 6 | 3231 | 0,0019 | 1,0000 |
| Celkem | 3231 | - | 1,0000 | - |

Medián

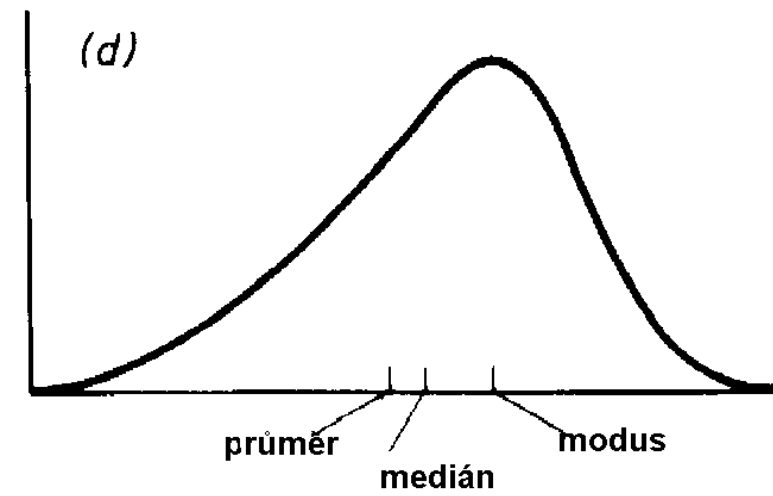
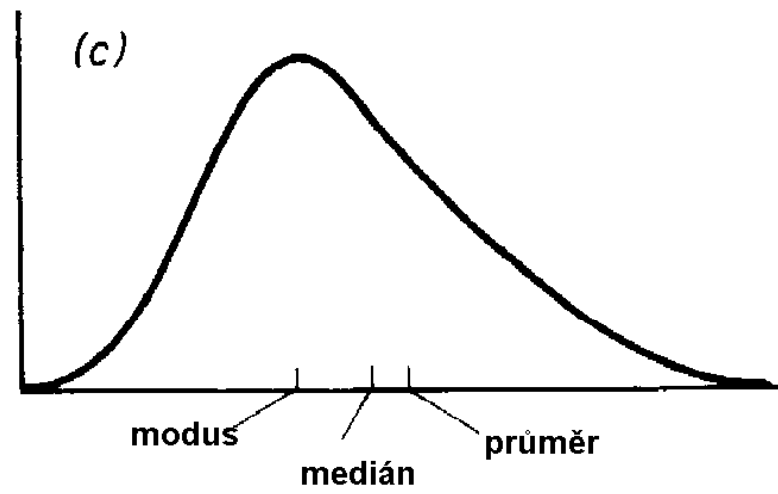
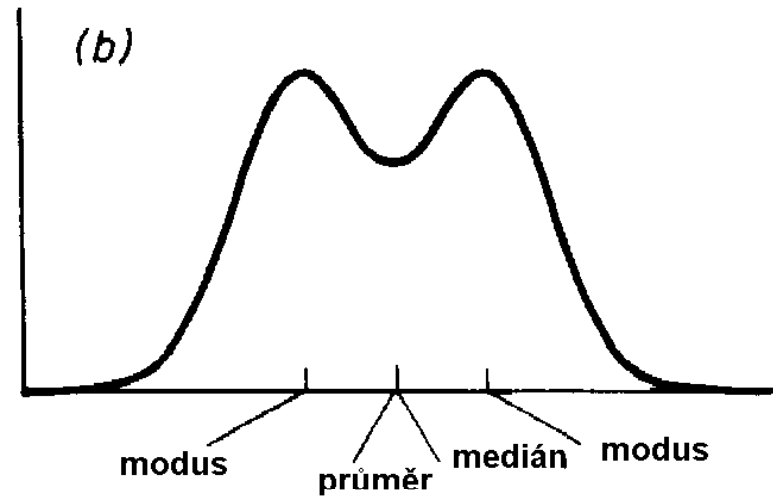
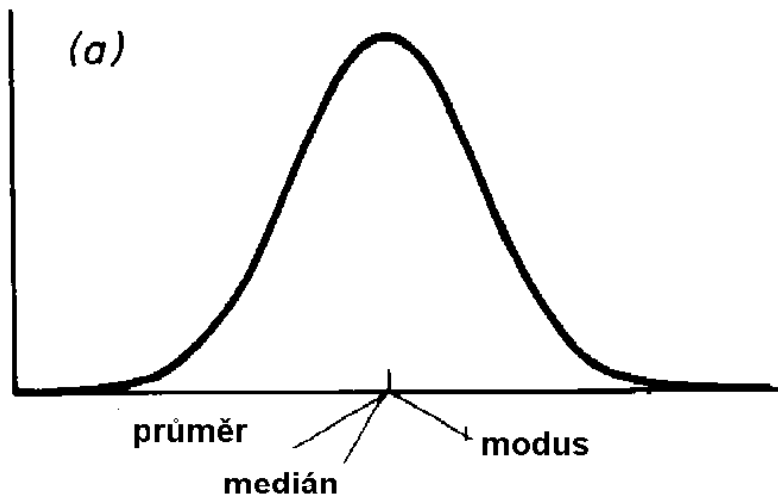
\tilde{x}

- Hodnota, která rozdělí pozorování na dvě stejně velké skupiny
- Kvantitativní a ordinální veličiny
- Měření 1: 73, 25, 15, 22, 50
- Měření 2: 61, 49, 35, 74, 53, 82
- Vzdělání: bez základního vzdělání, absolvent ZŠ, vyučen, vyučen s maturitou, vysokoškolák
- Není ovlivněn odlehlými pozorováními

Modus

\hat{x}

- Nejčastěji se vyskytující hodnota v souboru
- Kvalitativní (ordinální znaky)
- Kvantitativní znaky – modální interval => modus
- Vystižení nejtypičtější hodnoty
- Jablko, pomeranč, hruška, pomeranč, jablko, jablko, hruška



Geometrický průměr

$$\bar{x}_G = \sqrt[n]{x_1 x_2 \dots x_n} = \sqrt[n]{\prod_{i=1}^n x_i}$$

- Veličiny měřené na logaritmické stupnici
- Tempo růstu
- Průměrný koeficient růstu produkce jednoho podniku za celý rok
- V jednotlivých čtvrtletích byl koeficient růstu:
- 0,98; 1,02; 1,12; 1,05

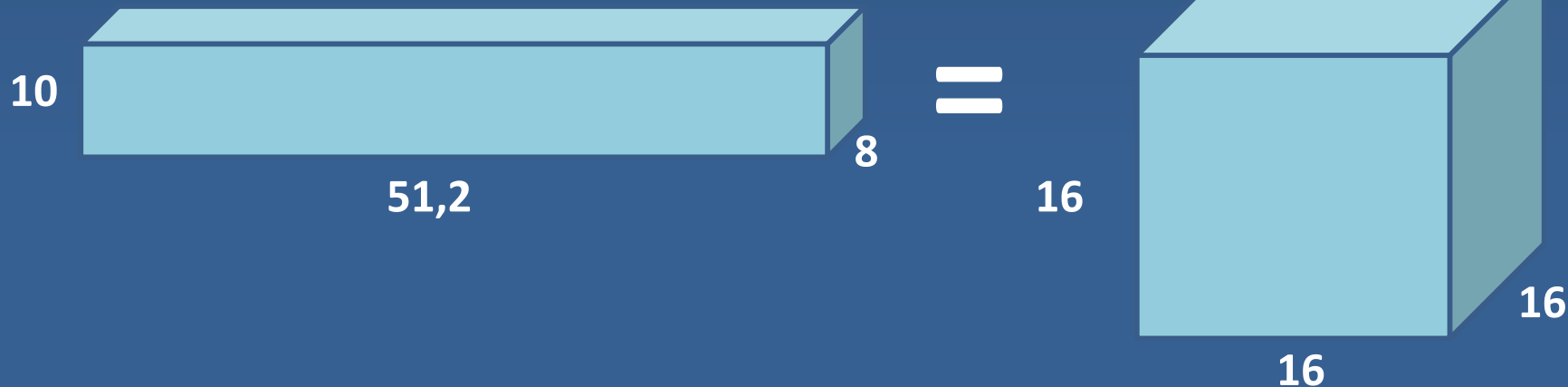
Geometrická interpretace geometrického průměru

- Geometrický průměr čísel 2 a 18



$$2 \times 18 = 6 \times 6$$

- Geometrický průměr čísel 10 a 51,2 a 8



$$10 \times 51,2 \times 8 = 16 \times 16 \times 16$$

Vliv odlehlých pozorování

- Počet hodin týdně strávených u televize
- 5, 7, 3, 38, 7
- 5, 7, 3, 38, 7, 200

Vliv odlehlých pozorování

- Platy ve dvou podnicích

| Podnik A | Podnik B | |
|--------------|--------------|---------------|
| 8000 | 7000 | |
| 9000 | 7500 | |
| 11000 | 8000 | |
| 12000 | 8500 | Medián |
| 15000 | 11000 | |
| 18000 | 18000 | |
| 20000 | 39000 | |
| | | |
| 13286 | 14143 | Průměr |

Průměrná x mediánová mzda v ČR

3/2019:

Průměrná: 33 840 Kč

Mediánová: 29 247 Kč

Kvantitativní a kvalitativní znaky

Data poměrová
(výška, hmotnost)



Data intervalová
(teplota ve °C)



Data ordinální
(závažnost onemocnění)



Data nominální
(pohlaví, typ operace)

Průměr

Medián

Modus