

3. Charakteristiky polohy

Statistické charakteristiky: čísla, která podávají stručnou a souhrnnou informaci o souboru

- úplnou statistickou informaci o znaku x podává jeho rozdělení četností
 - nejstručnější informaci pak podává jediné číslo charakterizující polohu znaku na číselné ose, popř. doplněné dalším číslem, které charakterizuje variabilitu (proměnlivost) hodnot znaku
- V tomto článku budeme vyšetřovat statistický soubor na základě pouze jednoho kvantitativního znaku a budeme se zabývat tzv. charakteristikami polohy
- Charakteristiku polohy chápeme jako číslo, kolem něhož jednotlivé hodnoty znaku kolísají.**

ARITMETICKÝ PRŮMĚR – zn. $\bar{x}$

- nejčastěji používaná charakteristika polohy
- má smysl jako charakteristika polohy, když jsou odchylky naměřených hodnot nahodilé a v souboru se nevyskytují extrémně nízké nebo vysoké hodnoty (pokud ne, pak použít další charakteristiky)
- **podíl součtu hodnot znaku u všech jednotek souboru a rozsahu souboru n** (součet hodnot znaku zjištěných u všech jednotek souboru, dělený počtem všech jednotek souboru)

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

Příklady

- ① Při výzkumu byly zjištěny tyto známky z fyziky: 3, 2, 1, 3, 2, 2, 3, 3, 4, 3, 3, 3, 3, 3, 3, 3, 1. Určete průměr známek z fyziky.

$$\bar{x} = \frac{3 + 2 + 1 + 3 + 2 + 2 + 3 + 3 + 4 + 3 + 3 + 3 + 3 + 3 + 3 + 3 + 1}{19} = \frac{51}{19} \doteq 2,68$$

- pro větší počet čísel výpočet zdlouhavý $\Rightarrow$ pro usnadnění lze využít tabulku četností

- ② Vypočti aritmetický průměr známek z fyziky z příkladu 1 s využitím tabulky četností jednotlivých hodnot.

známka x_j^*	1	2	3	4	5
četnost n_j	2	3	13	1	0

$$\bar{x} = \frac{2 \cdot 1 + 3 \cdot 2 + 13 \cdot 3 + 1 \cdot 4 + 0 \cdot 5}{19} = \frac{51}{19} \doteq 2,68$$

- **vzorec pro aritmetický průměr z tabulky rozdělení četností** (musíme každou hodnotu x_j^* násobit její četností – obdoba váženého aritmetického průměru)

$$\bar{x} = \frac{n_1 x_1^* + n_2 x_2^* + \dots + n_r x_r^*}{n} = \frac{1}{n} \sum_{j=1}^r n_j x_j^*$$

- ③ Určete průměrnou výšku postavy z tabulky rozdělení četností, $n = 194$

x_j^*	160	165	170	175	180	185	190
n_j	9	20	30	82	35	14	4

$$\bar{x} = \frac{9 \cdot 160 + 20 \cdot 165 + 30 \cdot 170 + 82 \cdot 175 + 35 \cdot 180 + 14 \cdot 185 + 4 \cdot 190}{194} \doteq 174,4$$

- ④ Průměrný roční čistý příjem v Kč byl zjištěn u všech 20 zaměstnanců podniku. Výsledky šetření jsou zapsány v následující tabulce rozdělení četností. Určete aritmetický průměr z tabulky rozdělení četností.

Roční příjem	151000–175000	176000–200000	201000–225000	226000–250000	251000–275000
Střední intervalů	163 000	188 000	213 000	238 000	263 000
Četnost n_j	4	5	8	2	1

$$\bar{x} = \frac{4 \cdot 163\,000 + 5 \cdot 188\,000 + 8 \cdot 213\,000 + 2 \cdot 238\,000 + 1 \cdot 263\,000}{20} = 201\,750$$

$$[\text{nebo } \bar{x} = \frac{1}{20} (4 \cdot 163\,000 + 5 \cdot 188\,000 + 8 \cdot 213\,000 + 2 \cdot 238\,000 + 1 \cdot 263\,000) = 201\,750]$$

– **metoda vhodně zvoleného začátku**

- jsou-li hodnoty znaku x velká čísla, lze výpočet hodnoty aritmetického průměru ještě zjednodušit využitím tzv. **metody vhodně zvoleného počátku**

- hodnoty znaku x zapíšeme ve tvaru $x = a + bu$, kde a, b jsou konstanty obvykle volíme konstanty a, b takto:
 a ... nejmenší hodnota, b ... šířka intervalů, lze i násobky
- **aritmetický průměr** potom je

$$\bar{x} = a + b \cdot \bar{u} = a + b \cdot \frac{1}{n} \sum_{i=1}^n u_i$$

- ⑤ Vypočítejte ještě jednou průměrnou výšku s využitím metody vhodně zvoleného počátku.

x_j^*	160	165	170	175	180	185	190
n_j	9	20	30	82	35	14	4
u_i	0	1	2	3	4	5	6

volíme $a = 160$, $b = 5$, tj. $x = 160 + 5u$

$$\bar{u} = \frac{0 \cdot 9 + 1 \cdot 20 + 2 \cdot 30 + 3 \cdot 82 + 4 \cdot 35 + 5 \cdot 14 + 6 \cdot 4}{194} = \frac{560}{194} \doteq 2,887$$

$$\bar{x} = a + b\bar{u} = 160 + 5 \cdot 2,887 = 174,4$$

- ⑥ Metodou vhodně zvoleného počátku vypočítejte průměrný roční hrubý příjem zaměstnanců v tomto podniku, pracujte se střední intervalů.

Roční příjem	151000–175000	176000–200000	201000–225000	226000–250000	251000–275000
Střední intervalů	163 000	188 000	213 000	238 000	263 000
Četnost n_j	4	5	8	2	1
a	163 000	163 000	163 000	163 000	163 000
$x_i^* - a$	0	25 000	50 000	75 000	100 000
b	25 000	25 000	25 000	25 000	25 000
u_i	0	1	2	3	4

$$\begin{aligned} \bar{x} &= a + b\bar{u} = a + b \cdot \frac{1}{n} \sum_{i=1}^n u_i = 163\,000 + 25\,000 \cdot \frac{(0 \cdot 4 + 5 \cdot 1 + 8 \cdot 2 + 2 \cdot 3 + 1 \cdot 4)}{20} = \\ &= 163\,000 + 25\,000 \cdot \frac{31}{20} = 163\,000 + 38\,750 = 201\,750 \end{aligned}$$

- vážený aritmetický průměr

- využíváme, pokud se soubor skládá z r dílčích souborů, v nichž známe počty jednotek v_j a průměry u_j a chceme vypočítat aritmetický průměr celého souboru

např. průměr prospěchu celé školy, známe-li průměry prospěchu jednotlivých tříd a počty žáků v jednotlivých třídách

- vážený průměr čísel $u_1, u_2, \dots, u_r$ s váhami $v_1, v_2, \dots, v_r$ ($v_j > 0$)

$$\bar{x}_v = \frac{\sum_{j=1}^r u_j v_j}{\sum_{j=1}^r v_j} = \frac{u_1 v_1 + u_2 v_2 + \dots + u_r v_r}{v_1 + v_2 + \dots + v_r}$$

- ⑦ Ve škole jsou čtyři třídy v ročníku označené A, B, C, D, počty žáků a průměrné známky jsou uvedeny v tabulce. Určete průměrnou známku z matematiky ze všech čtyř tříd dohromady.

třída	A	B	C	D
průměrná známka z M	2,21	1,82	2,33	2,11
počet žáků	28	24	32	30

$$\bar{x}_v = \frac{2,21 \cdot 28 + 1,82 \cdot 24 + 2,33 \cdot 32 + 2,11 \cdot 30}{28 + 24 + 32 + 30} = \frac{243,42}{114} \doteq 2,14$$

- ⑧ Ve škole jsou čtyři třídy v ročníku označené A, B, C, D, počty žáků a průměrné známky jsou uvedeny v tabulce mimo průměr třídy B. Určete průměrnou známku z matematiky třídy B, je-li průměr všech čtyř tříd dohromady.

třída	A	B	C	D
průměrná známka z M	1,92	$\bar{x}_B$	1,82	2,15
počet žáků	29	24	22	25

$$1,96 = \frac{1,92 \cdot 29 + \bar{x}_B \cdot 24 + 1,82 \cdot 22 + 2,15 \cdot 27}{29 + 24 + 22 + 25}$$

$$196 = 149,47 + 24\bar{x}_B$$

$$24\bar{x}_B = 46,53$$

$$\bar{x}_B \doteq 1,94$$

- ⑨ V tabulce je uvedena hustota obyvatel na 1 km² a celková rozloha v km² pěti středoevropských států. Jaká je průměrná hustota obyvatel v této části Evropy?

	Polsko	ČR	Slovensko	Rakousko	Maďarsko
hustota	124	131	110	97	110
rozloha	312 700	78 900	49 000	84 900	93 000

$$\begin{aligned}\bar{x}_v &= \frac{124 \cdot 312\,700 + 131 \cdot 78\,900 + 110 \cdot 49\,000 + 97 \cdot 84\,900 + 110 \cdot 93\,000}{312\,700 + 78\,900 + 49\,000 + 84\,900 + 93\,000} = \\ &= \frac{72\,966\,000}{618\,500} \doteq 118 \text{ obyv./km}^2\end{aligned}$$

- ⑩ V okrese jsou 3 zemědělská družstva a soukromě hospodařící rolníci. V tabulce jsou hektarové výnosy pšenice v t/ha a její sklizňové plochy v ha u všech čtyř subjektů. Jaký byl průměrný hektarový výnos pšenice v okrese?

	ZD1	ZD2	ZD3	soukromníci
výnos	4,9	5,2	4,8	5,4
plocha	640	560	1 100	400

$$\bar{x}_v = \frac{4,9 \cdot 640 + 5,2 \cdot 560 + 4,8 \cdot 1\,100 + 5,4 \cdot 400}{640 + 560 + 1\,100 + 400} = \frac{13\,488}{2\,700} \doteq 5,00 \text{ t/ha}$$

Doposud jsme uvažovali, že sledovaný znak existuje typická hodnota a kolem ní náhodné odchylky. Existují i jiné typy dat, kdy odchylky nejsou nahodilé, ale systematické, např. v časových řadách, kde data vykazují určitý trend, vývoj v čase, aritmetický průměr pak ztrácí smysl – lepším ukazatelem než průměr bývá průměrný přírůstek (úbytek), průměrné tempo růstu vyšetřovaného znaku za jedno období.

- ii) V tabulce je uveden HDP ČR na obyvatele v letech 2000 až 2009 v tisících Kč. Nemá smysl počítat průměrný HDP na obyvatele mezi lety 2000–2009, HDP se téměř neustále zvyšoval, hodnoty na počátku období jsou podstatně nižší než na konci.

Rok	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
HDP (x_i)	213	230	242	253	276	292	314	342	354	346
Roční přírůstek (y_i)		17	12	11	23	16	22	28	12	-8
Roční tempo růstu (z_i)		1,080	1,052	1,045	1,091	1,058	1,075	1,089	1,035	0,977

- a) Doplně do tabulky roční přírůstky. Ve kterém roce rostl HDP nejrychleji?

$$\text{Roční přírůstek } y_i = x_i - x_{i-1}$$

Nejrychleji rostl HDP v roce 2007, kdy oproti předchozímu roku vzrostl o 28 tisíc Kč na obyvatele.

- b) Urči průměrný přírůstek HDP ČR mezi lety 2000 až 2009

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{17 + 12 + 11 + 23 + 16 + 22 + 28 + 12 - 8}{9} = \frac{133}{9} = 14,8$$

$$\text{nebo lze vypočítat i jinak } \bar{y} = \frac{x_n - x_0}{n} = \frac{346 - 213}{9} = 14,8$$

V roce 2002 a 2008 byl průměrný přírůstek stejný, ale přesto byl rok 2002 zřejmě úspěšnější, protože HDP se zvyšovalo z nižší úrovně $\Rightarrow$ lépe charakterizuje tzv. roční tempo růstu.

- c) Určete roční tempo růstu a doplňte do tabulky. Ve kterém roce rostl HDP nejrychleji?

$$\text{Roční tempo růstu } \bar{z} = \frac{x_i}{x_{i-1}}$$

Nejrychleji rostl HDP v roce 2004, kdy oproti předchozímu roku vzrostl 1,091krát.

- d) Určete průměrné tempo růstu.

Pokud chceme určit hodnotu HDP po dvou letech (v roce 2002): $242 = 213 \cdot 1,08 \cdot 1,052 \Rightarrow$ ve vzorci hodnoty násobíme $\Rightarrow$ místo aritmetického průměru použijeme průměr geometrický

$$\bar{z}_G = \sqrt[n]{z_1 \cdot z_2 \cdot \dots \cdot z_n} = \sqrt[9]{1,080 \cdot 1,052 \cdot \dots \cdot 0,977} = 1,055$$

$$\text{nebo lze vypočítat i jinak } \bar{z}_G = \sqrt[n]{\frac{x_n}{x_0}} = \sqrt[9]{\frac{346}{213}} = 1,055$$

- přírůstky za jednotlivá období

$$y_i = x_i - x_{i-1} \quad (y_1 = x_1 - x_0, y_2 = x_2 - x_1, y_3 = x_3 - x_2, \dots, y_n = x_n - x_{n-1})$$

PRŮMĚRNÝ PŘÍRŮSTEK - zn. $\bar{y}$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{x_n - x_0}{n}$$

- k výpočtu průměrného přírůstku není třeba určovat přírůstky za jednotlivá období, stačí celkový přírůstek za n období dělit jejich počtem

$$\left[\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{(x_1 - x_0) + (x_2 - x_1) + (x_3 - x_2) + \dots + (x_{n-1} - x_{n-2}) + (x_n - x_{n-1})}{n} = \frac{(-x_0 + x_n)}{n} = \frac{(x_n - x_0)}{n} \right]$$

Průměrné tempo růstu za jedno období – zajímá nás např. v národohospodářských časových řadách

– za průměr se zde nebere aritmetický průměr, ale tzv. geometrický průměr

– jde o průměr podílů hodnot za dvě po sobě následující období, tedy podílů

$$z_1 = \frac{x_1}{x_0}, z_2 = \frac{x_2}{x_1}, \dots, z_n = \frac{x_n}{x_{n-1}}, \text{ (většinou se udávají v \%, pak } z_n = \frac{x_n}{x_{n-1}} \cdot 100 \%)$$

GEOMETRICKÝ PRŮMĚR – zn. $\bar{z}_G$

$$\bar{z}_G = \sqrt[n]{z_1 \cdot z_2 \cdot \dots \cdot z_n} = \sqrt[n]{\frac{x_n}{x_0}} \quad \text{– zavádí se jen pro kladná čísla}$$

Příklady

- ⑫ V pěti po sobě jdoucích letech jsou hodnoty růstu 101,3; 108,5; 100,6; 104,2; 102,1 v procentech. Jaké je průměrné roční tempo růstu?

$$\bar{z}_G = \sqrt[5]{z_1 \cdot z_2 \cdot \dots \cdot z_n} = \sqrt[5]{101,3 \cdot 108,5 \cdot 100,6 \cdot 104,2 \cdot 102,1} \doteq 103,3 \%$$

- ⑬ V rozvojové zemi došlo během určitého roku k prudkému vzrůstu inflace. V tabulce je pro každý měsíc uveden index spotřebitelských cen v procentech proti minulému měsíci (např. v únoru byly ceny v průměru o 4 % vyšší než v lednu, ceny v prosinci byly o 80 % vyšší než v listopadu). Jaké bylo průměrné měsíční tempo růstu cen? Kolikrát byly ceny v prosinci toho roku vyšší ve srovnání s prosincem minulého roku?

I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
102	104	104	104	110	120	130	140	150	160	170	180

– průměrné měsíční **tempo** růstu cen

$$\bar{z} = \sqrt[12]{102 \cdot 104 \cdot 104 \cdot 104 \cdot 110 \cdot 120 \cdot 130 \cdot 140 \cdot 150 \cdot 160 \cdot 170 \cdot 180} = 128,5 \%$$

– ceny byly v prosinci ve srovnání s prosincem minulého roku vyšší

$$p = 1,02 \cdot 1,04 \cdot 1,04 \cdot 1,04 \cdot 1,10 \cdot 1,20 \cdot 1,30 \cdot 1,30 \cdot 1,40 \cdot 1,50 \cdot 1,60 \cdot 1,70 \cdot 1,80 = 20,25\text{krát}$$

HARMONICKÝ PRŮMĚR – zn. $\bar{x}_H$

– převrácená hodnota aritmetického průměru převrácených hodnot

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

Příklady

- ⑭ Součástka se vyrábí na dvou automatech. Starší automat vyrobí 1 kus každých 6 minut, nový každé 3 minuty. Jak dlouho trvá v průměru výroba jedné součástky?

– starší za 1 hodinu 10 ks, nový 20 ks za 1 hodinu, v průměru každý z nich (10 + 20) : 2 = 15 ks za hodinu, tedy 1 ks za (60 min : 15) = 4 minuty

– lze přímo jako harmonický průměr hodnot $x_1 = 6$ a $x_2 = 3$:

$$\bar{x} = \frac{1}{\frac{1}{n} \left(\frac{1}{x_1} + \dots + \frac{1}{x_n} \right)} = \frac{1}{\frac{1}{2} \left(\frac{1}{6} + \frac{1}{3} \right)} = \frac{1}{\frac{1}{2} \cdot \frac{3}{6}} = \frac{1}{\frac{1}{4}} = 4$$

- ⑮ Stezkou, která vede na vrchol hory, vystupuje turista rychlostí 2,5 km/h, sestupuje rychlostí 5 km/h. Jakou průměrnou rychlostí jde?

- průměrná rychlost: $v_p = \frac{\text{celková dráha}}{\text{celková doba}} = \frac{s}{t}$

$$v_p = \frac{s'}{t} = \frac{s_1 + s_2}{t_1 + t_2} = \frac{2s}{\frac{s}{v_1} + \frac{s}{v_2}} = \frac{2s}{\frac{sv_2 + sv_1}{v_1 v_2}} = \frac{2s}{\frac{s(v_2 + v_1)}{v_1 v_2}} = \frac{2v_1 v_2}{v_1 + v_2}$$

$$\left[\begin{array}{l} s' = s_1 + s_2 = s + s = 2s \\ \text{tam: } s_1 = v_1 t_1 \rightarrow t_1 = \frac{s_1}{v_1} = \frac{s}{v_1} \quad \text{zpátky: } s_2 = v_2 t_2 \rightarrow t_2 = \frac{s_2}{v_2} = \frac{s}{v_2} \end{array} \right]$$

$$v_p = \frac{2v_1 v_2}{v_1 + v_2} = \frac{2 \cdot 2,5 \cdot 5}{2,5 + 5} = \frac{25}{7,5} = \frac{10}{3} \doteq 3,3 \frac{\text{km}}{\text{h}}$$

- lze i pomocí harmonického průměru:

$$\bar{x}_H = \frac{1}{\frac{1}{n} \left(\frac{1}{x_1} + \dots + \frac{1}{x_n} \right)} = \frac{1}{\frac{1}{2} \left(\frac{1}{v_1} + \frac{1}{v_2} \right)} = \frac{1}{\frac{1}{2} \left(\frac{1}{2,5} + \frac{1}{5} \right)} = \frac{2}{\frac{5 + 2,5}{2,5 \cdot 5}} = \frac{2 \cdot 2,5 \cdot 5}{7,5} = \frac{10}{3} \doteq 3,3 \text{ km/h}$$

4. Doplnující charakteristiky polohy – modus a medián

Příklady

- ① Ve třídě s 31 studenty byla zjišťována výše jejich kapesného na měsíc. Výsledky šetření jsou zpracovány v následující tabulce rozdělení četností.

Výše kapesného	500	1 000	2 000	5 000
Četnost studentů	15	12	3	1

Určete

- a) průměrnou výši kapesného na týden,
 b) výši kapesného s největší četností (modus)
 c) prostřední hodnotu znaku (medián) kapesného ve třídě.
 Porovnejte tyto charakteristiky polohy.

$$\text{a) } \bar{x} = \frac{15 \cdot 500 + 12 \cdot 1\,000 + 3 \cdot 2\,000 + 1 \cdot 5\,000}{31} = \frac{30\,500}{31} \doteq 984 \text{ Kč}$$

$$\text{b) výše kapesného s největší četností } \text{Mod}(x) = 500 \text{ Kč}$$

$$\text{c) prostřední hodnota výše kapesného } \text{Med}(x) = x_{(16)} = 1000 \text{ Kč}$$

Protože jedna hodnota znaku je v tomto souboru výrazně odlišná, je jako doplňková charakteristiky polohy vhodnější medián (než modus – je výrazně nižší než $\bar{x}$ i Med).

- ② a) Souborem jsou pracovní úrazy za určité období v podniku s nepřetržitým provozem. Znakem je denní doba, kdy k pracovnímu úrazu došlo, např. 24 hodin rozdělíme na intervaly po 2 hodinách.

Modus je tedy ta dvouhodinová část denní doby, v níž nejčastěji dochází k úrazům. Jeho zjištění má zřejmě význam při hledání příčin úrazů a při jejich prevenci.

- b) Sledujeme počet automobilů, které projedou daným místem v různou hodinu dne, nebo počet chodců, kteří přecházejí danou křižovatkou.

Modus zde představuje hodinu nejživějšího provozu.

MODUS znaku x – zn. $\text{Mod}(x)$

- je **hodnota x , která má v souboru největší četnost**
- používáme jako doplňující charakteristiku souboru tehdy, máme-li sestavenou tabulku rozdělení četností s velkým rozsahem hodnot nebo má-li specifický význam

MEDIÁN znaku x – zn. $\text{Med}(x)$

- je **prostřední hodnota znaku**, jsou-li **hodnoty $x_1, x_2, \dots, x_n$ uspořádány podle velikosti**
- používáme jako doplňkovou charakteristiku souboru tehdy, když jsou v souboru zastoupeny prvky s hodnotami znaku mimořádně odlišnými oproti ostatním hodnotám
- je-li $x_1 \leq x_2 \leq \dots \leq x_n$, pak

1. pro n **liché** je $\text{Med}(x) = x_{\frac{n+1}{2}}$

2. pro n **sudé** je $\text{Med}(x) = \frac{1}{2} \left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1} \right)$

③ Určete modus a medián

a) v příkladě 1 čl. 2 (rodinní příslušníci)

$$\text{Mod}(x) = 4$$

$$\text{Med}(x) = x_{(72)} = 3$$

b) v příkladě 2 čl. 2 (známka z testu)

$$\text{Mod}(x) = 3$$

$$\text{Med}(x) = \frac{1}{2} (x_{(16)} + x_{(17)}) = \frac{3 + 3}{2} = 3$$

c) v příkladě 3 čl. 2 (výška)

$$\text{Mod}(x) = 175$$

$$\text{Med}(x) = \frac{x_{(97)} + x_{(98)}}{2} = 175$$

④ Souborem je 20 členů družstva, znakem x jejich roční příjem (v tisících Kč). Určete průměrný roční příjem a medián.

roční příjem	80	90	100	110	120	890
četnost	1	6	6	5	1	1

$$- n = 20$$

– průměrný roční příjem

$$\bar{x} = \frac{1 \cdot 80 + 6 \cdot 90 + 6 \cdot 100 + 5 \cdot 110 + 1 \cdot 120 + 1 \cdot 890}{20} = \frac{2\,780}{20} = 139 \text{ (tisíc) Kč}$$

– medián (n sudé)

$$\text{Med}(x) = \frac{x_{(10)} + x_{(11)}}{2} = \frac{100 + 100}{2} = 100 \text{ (tisíc) Kč}$$

Medián je zde vhodnější charakteristikou, protože mimo jednoho člena mají všichni podstatně nižší příjem.