

9.3.2 Četnost, relativní četnost

Předpoklady: 9301

Pedagogická poznámka: Na tuto a následující hodiny je nutné zajistit papírky s kompletními daty výzkumu, aby žáci mohli zpracovávat data přímo v lavicích.

Tabulka prvních deseti známek z fyziky:

jednotka	1	2	3	4	5	6	7	8	9	10
známka	3	2	1	3	2	2	3	3	3	4

Celkem máme n jednotek s hodnotami všech zkoumaných znaků.

Hodnoty znaku pro jednotlivé jednotky značíme $x_1, x_2, \dots, x_i, \dots, x_n$.

Často může znak nabývat pouze určitého počtu hodnot \Rightarrow počet hodnot r .

Jednotlivé hodnoty značíme $x_1^*, x_2^*, \dots, x_j^*, \dots, x_r^*$ (pozor poslední index je r , protože různých hodnot je pouze r) \Rightarrow zápis hodnot můžeme ulehčit tím, že v tabulce uvedeme pouze možné hodnoty znaku a kolikrát se každá hodnota x_j^* vyskytla – tedy četnost n_j hodnoty x_j^* .

Pedagogická poznámka: Předchozí rozdělení je stálým zdrojem nedorozumění. Snažím se, aby žáci měli v paměti dva páry pojmů: počet jednotek – n , počet různých hodnot – r .

Př. 1: Sestav tabulku četností známek z fyziky (nakresli ji tak, aby měla tři řádky, poslední zatím zůstává volný).

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

\Rightarrow Získali jsme **rozdělení četností** (obsahuje ze statistického hlediska úplnou informaci).

Pedagogická poznámka: Doporučuji žákům, aby si tuto tabulku přepsali na samostatný papír a všechny další tabulky četností psali rovnou na tento papír místo do sešitu. Hodnoty pak budou mít v dalších hodinách lépe po ruce.

Př. 2: Najdi vztah, který musí platit pro četnosti libovolného znaku.

Když všechny četnosti sečteme, musíme dostat počet všech zkoumaných jednotek, tedy n .

$$\sum_{j=1}^r n_j = n_1 + n_2 + \dots + n_r = n.$$

Četnost jedničky z fyziky je jedna, ale tento samotný údaj neříká, jak častá tato známka je, roli hraje i to, kolik lidí test vyplňovalo (deset, jedna třída, celá škola) \Rightarrow **relativní četnost**

znaku $v_j = \frac{n_j}{n}$ (udává, jak častý je znak vzhledem k počtu všech zkoumaných jednotek).

Př. 3: Doplň tabulku s rozdělením četností známek z fyziky a najdi podmínku, kterou musí hodnoty četnosti splňovat.

známka x_j^*	1	2	3	4	5
četnost n_j	1	3	5	1	0
relativní četnost v_j	0,1	0,3	0,5	0,1	0

Součet relativních četností se rovná jedné: $\sum_{j=1}^r v_j = v_1 + v_2 + \dots + v_r = 1$.

Př. 4: Sestav tabulku četností a relativních četností všech známek z matematiky. Zkontroluj součty.

Tabulka četností známek z matematiky.

známka x_j^*	1	2	3	4	5
četnost n_j	1	5	10	3	0
relativní četnost v_j	0,053	0,263	0,526	0,158	0

Součet četností: $\sum_{j=1}^r n_j = 1 + 5 + 10 + 3 = 19$

Součet relativních četností: $\sum_{j=1}^r v_j = 0,053 + 0,263 + 0,526 + 0,158 = 1$

Některých znaky mohou mít příliš mnoho hodnot. Například výška (pokud bychom měřili na cm) by vyžadovala tabulku možná až se 100 sloupci \Rightarrow při vytváření tabulky sdružujeme hodnoty do intervalů.

Například výšky účastníků výzkumu můžeme sdružit do intervalů po 5 cm. Jaké intervaly zvolíme?

Zdánlivě nesmyslně: $(157,5;162,5), (162,5;167,5), (167,5;172,5), \dots$

Proč takhle?

Při sdružování do intervalů zaokrouhlujeme na středy intervalů, které v tomto případě vycházejí velmi hezky: $(157,5;162,5) \Rightarrow 160, (162,5;167,5) \Rightarrow 165, (167,5;172,5) \Rightarrow 170, \dots$ intervaly pak můžeme popsat jejich středy.

Př. 5: Sestav tabulku četností výšky, použij intervaly po 5 cm se středy 155, 160, 165,...

x_j^*	160	165	170	175	180	185	190
n_j	1	1	4	6	4	2	1

Tabulka z předchozího příkladu by mohla být zapsána také následujícími způsoby:

x_j^*	157,5- 162,5	162,5 – 167,5	167,5 – 172,5	172,5 – 177,5	177,5 – 182,5	182,5 – 187,5	187,5 – 192,5
---------	-----------------	------------------	------------------	------------------	------------------	------------------	------------------

n_j	1	1	4	6	4	2	1
-------	---	---	---	---	---	---	---

popis intervalů je takto příliš dlouhý a proto se zkracuje většinou takto:

x_j^*	157-162	162 – 167	167 – 172	172 – 177	177 – 182	182 – 187	187 – 192
n_j	1	1	4	6	4	2	1

což ne zcela odpovídá pravidlům pro zaokrouhlování. Navíc tento zápis sice dobře přibližuje, kam máme zapisovat hodnoty zaokrouhlené na cm, ale neříká příliš jasně, kam máme zařadit hodnoty 167,2 nebo ještě hůře 167,5.

Ještě nejasnější jsou zápisy tabulek při jiné volbě intervalu:

x_j^*	160-164	165 – 169	170 – 174	...
n_j	1	1	4	..

Striktně podle předchozího postupu jsou středy intervalů čísla 162, 167, 172, ... a samotné intervaly $(159,5;164,5)$, $(164,5;169,5)$, $(169,5;174,5)$, ..., což je rozdělení, které by nikdo sám od sebe nepoužil (přirozenější je určitě dělení na intervaly $(160;165)$, $(165;170)$, $(170;175)$, ... se středy v bodech 162,5; 167,5; 172,5; ...).

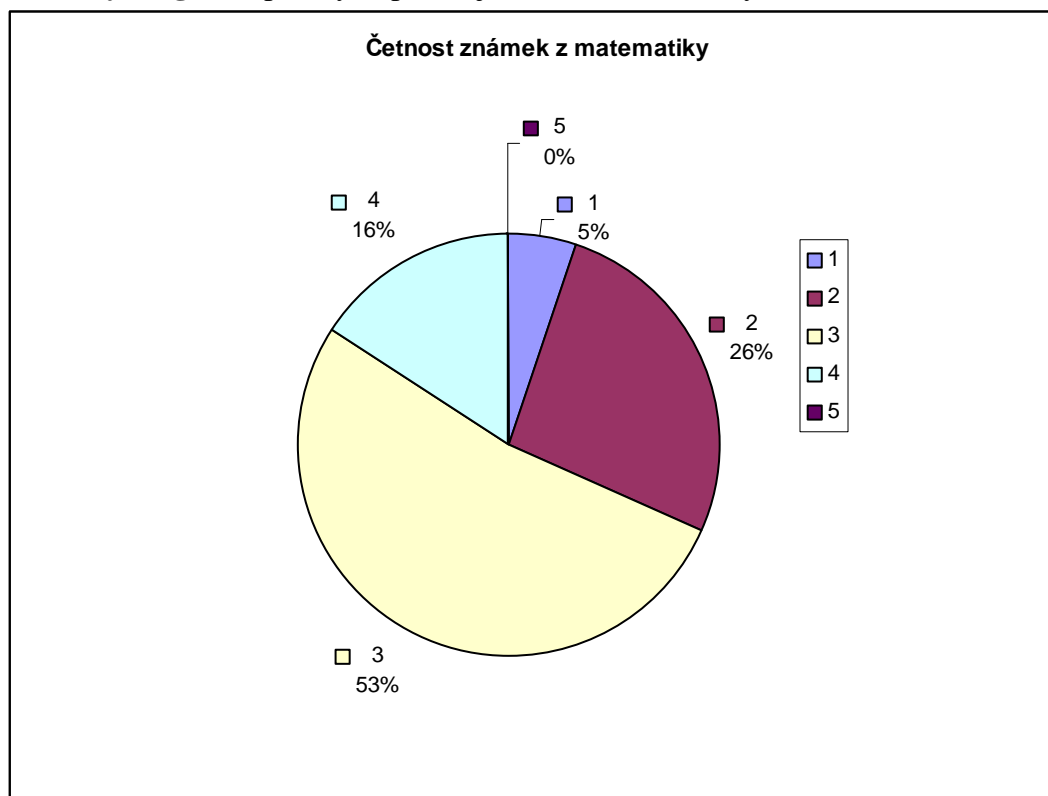
Z předchozího je vidět, že daleko přehlednější je popisovat tabulku četností pomocí středů.

Kvůli rychlejšímu pochopení dat, se tabulky často zobrazují do grafů.

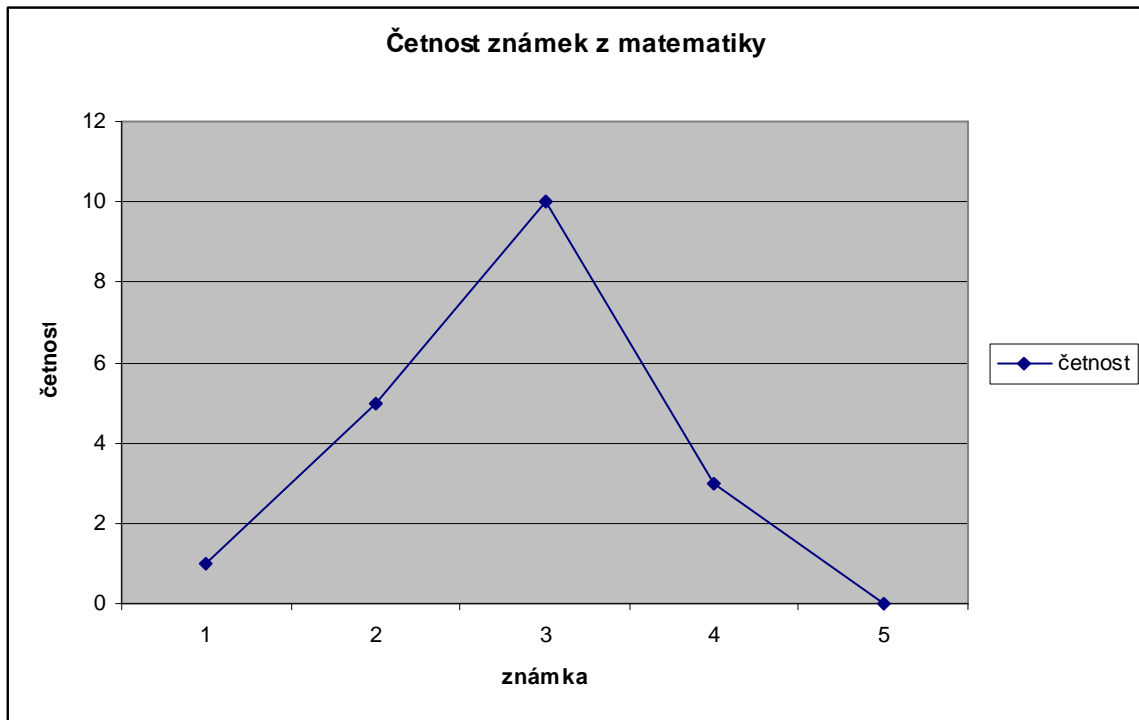
Tabulka četností známek z matematiky:

známka x_j^*	1	2	3	4	5
četnost n_j	1	5	10	3	0
relativní četnost v_j	0,053	0,263	0,526	0,158	0

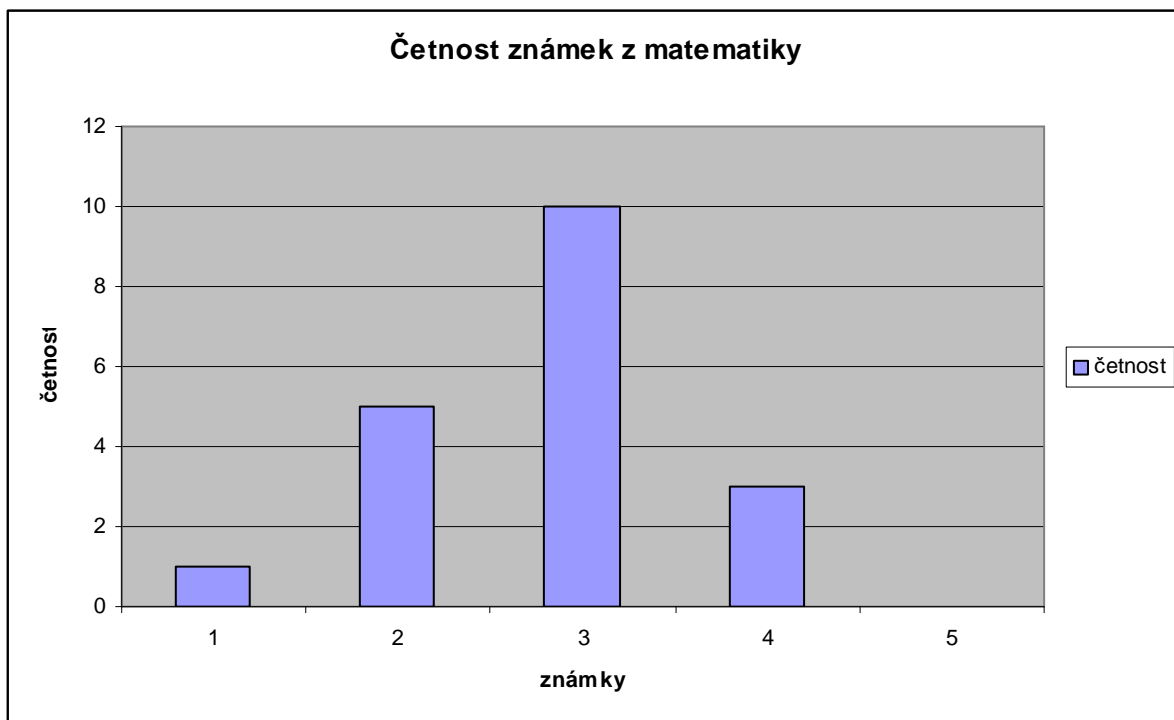
Kruhový diagram (plochy odpovídají relativní četnosti výsledků)



Spojnicový digram (výška bodů odpovídá jednotlivým četnostem)



Sloupcový digram (výška sloupců odpovídá jednotlivým četnostem)



Shrnutí: Úplnou statistickou informací zachycuje i tabulka četností (ke každé hodnotě je zapsáno, kolikrát se vyskytla).