

§14. Statistika

- Def:** *Statistický soubor* je množina všech objektů statistického pozorování
- Def:** *Statistická jednotka* prvek statistického souboru.
- Def:** *Rozsah souboru* n je počet prvků statistického souboru.
- Def:** *Statistický znak* je společná vlastnost statistických jednotek, zpravidla se značí x .
- Def:** *Hodnota znaku* = jednotlivé údaje znaku Příklady stat. znaků: Při sčítání obyvatelstva: věk, pohlaví, zaměstnání...
Hodnoty znaku mohou být vyjádřeny slovy nebo čísly (kvalitativní \times kvantitativní znak)
- Def:** *Četnost (absolutní četnost)* hodnoty znaku je počet statistických jednotek, které mají stejnou hodnotu znaku
Značí se n_j a platí $\sum_{j=1}^k n_j = n$.
- Def:** *Relativní četnost hodnoty znaku* je podíl absolutní četnosti a rozsahu souboru.
Značíme $\nu_j = \frac{n_j}{n}$ a platí $\sum_{j=1}^k \nu_j = 1$.
- Def:** *Rozdělení četnosti* – všechny různé hodnoty znaku a jim odpovídající četnosti uspořádáme do tabulky nebo znázorňujeme graficky.

- Spojnicový diagram (polygon četnosti)
- Sloupkový diagram (histogram)
- Kruhový diagram

- Def:** *Skupinové (intervalové) rozdělení četnosti* – blízké hodnoty znaku se sdružují do skupin tvořených obvykle intervaly. Hodnoty znaku, jež se dostali do téhož intervalu, lze potom reprezentovat jednou hodnotou – střed intervalu = třídní znak.

A) Charakteristiky statistického souboru

- Pozn:** Čísla, která podávají stručnou souhrn informací o uvažovaném statistickém souboru z různých hledisek.

B) Charakteristiky polohy

- Pozn:** Čísla, kolem nichž jednotlivé hodnoty znaku kolísají
- Def:** *Aritmetický průměr* hodnot x_1, x_2, \dots, x_n kvantitativního znaku x : $\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$
- Def:** *Vážený aritmetický průměr* statistického souboru, kde četnost hodnoty znaku x_i je n_i : $\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_k x_k}{n}$
- Def:** *Geometrický průměr* hodnot x_1, x_2, \dots, x_n kvantitativního znaku x : $\bar{x}_G = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$
- Def:** *Harmonický průměr* hodnot x_1, x_2, \dots, x_n kvantitativního znaku x : $\bar{x}_H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}$

V.14.1.: AG-nerovnost:

$$\bar{x} \geq \bar{x}_G$$

Def: *Modus* hodnot x_1, x_2, \dots, x_n znaku x je hodnota, která má největší četnost. Značí se $\text{Mod}(x)$.

Modus se užívá k odhadu střední hodnoty znaku souboru, nejčastěji tedy, máme-li sestavenou tabulku rozdělení četnosti pro statistický soubor s velkým rozsahem.

Def: *Medián* hodnot x_1, x_2, \dots, x_n znaku x statistického souboru, v němž jsou prvky uspořádány dle velikosti hodnot sledovaného znaku je prostřední hodnota:

- V souboru s lichým rozsahem se rovná prostřednímu členu.
- V souboru se sudým počtem znaků se rovná aritmetickému průměru dvou prostředních znaků s indexy $\frac{n}{2}$ a $\frac{n}{2} + 1$.

Př: Ve třídě 1.A je 16 děvčat: Údaje o výšce udává následující tabulka:

Výška	160 – 164	165 – 169	170 – 174	175 – 179	180 – 184
Střed intervalu	162	167	172	177	182
Počet	2	5	4	3	2

Průměrná výška: 171 cm

Modus: 167 cm

Medián: $\frac{172+172}{2} = 172$ cm

Př: V testu při zkoušce dostalo 15 studentů známku 1, dalších 35 2, 30 3, 15 4 a zbylých 5 studentů dostalo známku 5.

Vypočítejte průměrnou známku, modus a medián.

Průměr: 2.6

Medián: 2.5

Modus: 2

Př: Vypočítejte pro $n = 4, 8, 12, 16$, že v právě $\frac{1}{2}$ hodů padne pana.

Celkem možností hodů: 2^n Z toho padne pana polovině právě v $\left(\frac{n}{2}\right)$ případech. Celková pravděpodobnost tedy je:

$$\frac{n!}{2^n \cdot \frac{n!}{2} \cdot \frac{n!}{2}}$$

Číselně:

$$n = 4 \Rightarrow 0.375$$

$$n = 8 \Rightarrow 0.2734375$$

$$n = 12 \Rightarrow 0.2255859375$$

$$n = 16 \Rightarrow 0.196380615234375$$

Př: Házeme 5 kostkami, určete pravděpodobnost toho, že na 3 z nich padnou stejná čísla:

Počet všech možností: 6^5

Šance, že padne x na alespoň 3 kostkách: $\binom{3}{5} \cdot 5^2 + \binom{4}{5} \cdot 5 + \binom{5}{5} = 10 + 5 + 1 \dots$

Př: Co je pravděpodobnější? Nad rovnocenným partnerem zvítězím v 3 z 4 nebo z 5 z 8:

$$3 \text{ z } 4: \frac{\binom{3}{4}}{2^4}$$

Z města X do Y se lze dostat 2 cestami dle mapy: Jaká je pravděpodobnost, že dojede do Y :

$$\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} + \frac{1}{3} = 512$$

Př: Maturant chce pokračovat na PřF a vybere si jeden z oborů M,F,Ch,Bi s pravděpodobností dle konzultanta 0.4, 0.3, 0.2, 0.1. Konzultant neví, co si zvolil, ale dozvěděl se, že má výbornou. Pravděpodobnosti na výbornou jsou 0.1, 0.2, 0.3, 0.4. Vyčíslete pravděpodobnosti:

Pravděpodobnosti na výběr a zároveň 1 jsou: 0.04, 0.06, 0.06, 0.04. Celkem tedy 0.128.
Doplňím na celek: 0.3125, 0.46875, 0.46875, 0.3125.

Př: Při vyšetřování pacienta je podezření na 3 navzájem se vylučující choroby s pravděpodobnostmi 0.3 0.5 0.2. Laboratorní zkouška dává pozitivní u 15%, 30%, 20% nemocných danou nemocí. Jaké jsou pravděpodobnosti nemocí po kladné lab. zkoušce?

Př: 2× kostkou. Jaká je pravděpodobnost, že padle alespoň 10 ok, za předpokladu, že (právě při 1 z hodů, alespoň při 1 z hodů, při 1. hodu) padne 6.

Právě při jednom hodu: $\frac{2}{5}$

Alespoň při jednom hodu: $\frac{5}{11}$

Při prvním hodu: $\frac{3}{6} = \frac{1}{2}$