

## METODY V GEOGRAFII

Mgr. Darina MÍSAŘOVÁ, Ph.D.

Sylabus přednášky 4: Popisné statistické metody, Základní statistické charakteristiky

*Sylabus slouží jako přehled základních pojmů zmiňovaných na přednášce. Není dostačující pro úspěšné zvládnutí zkoušky z Metod v geografii. Sylabus je nezbytné doplnit informacemi z přednášky.*

### POPISNÉ STATISTICKÉ METODY

Deskriptivní metody:

1. přehledné vyjádření výsledků pomocí četnostních tabulek a grafů
2. Vypočty a grafické znázornění základních popisných statistických charakteristik

#### Rozdělení četností

- Statistické údaje jednotlivých statistických souborů pro další zpracování uspořádáváme
- U jednotek statistického souboru můžeme na základě kvantitativních hodnot zjišťovat jejich četnost – frekvenci.
- Četnost - počet prvků se stejnou hodnotou statistického znaku
- Používáme ho pro nespojitě znaky a při malém počtu variant

#### Skupinové rozdělení četností

- Pro spojité znaky udáváme počet prvků s hodnotami znaku patřícími do určitého intervalu (třídy).
- Jednotky statistického souboru roztřídíme podle velikosti do několika intervalů.
- Dolní a horní hranice (mez) intervalu udává, jakou nejmenší a největší hodnotu znaku do daného intervalu zařadíme.
- Délka či **šířka intervalu** je kladný rozdíl dvou po sobě následujících dolních (horních) mezí.
- Krajní interval může být otevřený (neuzavřený).
- U skupinového rozdělení četností zastupuje hodnoty znaku **střed intervalu** ( $x_s$ ).

Zásady pro stanovení hranic intervalů:

#### Četnosti

- absolutní
- relativní
- kumulované

#### Grafické znázornění rozdělení četností

**Histogram** – typ sloupcového diagramu

**Polygon** – spojnicový diagram

**Čára kumulovaných četností** – součtová čára, četnosti, vynášíme k horní hranici intervalu

## **Popisná statistika**

K čemu je to dobré?

Základní statistické charakteristiky „popisují“ statistický soubor

- a) charakteristiky úrovně – tzv. střední hodnoty
- b) charakteristiky variability
- c) charakteristiky asymetrie
- d) charakteristiky špičatosti

Výchozí data – způsob vypočtu

- z reálných hodnot
- ze skupinového rozdělení četností (reálné hodnoty seskupené do intervalů)

## **Charakteristiky úrovně**

Střední hodnoty umožňují porovnávání souborů

Jedná se o čísla, která reprezentují jednotlivé hodnoty statistického znaku, udávají polohu, charakterizují obecnou velikost jevu.

### **Střední hodnoty**

aritmetický průměr (+ vážený aritm. průměr, geometrický průměr, harmonický průměr)

modus

aritmetický střed

medián a kvantily

geografický medián

### **Aritmetický průměr**

nejčastěji používaná st. charakteristika

typický a netypický průměr

(jedno a více vrcholová rozdělení četností)

typický aritm. průměr – jednovrcholové rozdělení četností + blízký nejčetnější hodnotě

Vlastnosti aritmetického průměru

### **Vážený aritmetický průměr**

Příklady použití:

## **Modus**

modus - nejčtenější hodnota kvantitativního znaku ve studovaném souboru  
významný především u souboru nespojitých veličin

modální interval – interval zahrnující největší počet jednotek, závisí však na stanovení hranic intervalů

rozdělení s více mody – polymodální rozdělení

## **Aritmetický střed**

Aritm. střed je polovina součtu min. a max. hodnoty znaku v souboru  
pokud soubor obsahuje extrémní hodnoty, je aritmetický střed značně zkreslující  
charakteristika

## **Medián**

Medián – tzv. prostřední hodnota,

je to prvek řady uspořádané v neklesajícím pořadí ( od nejm. po největší), který ji dělí na dvě poloviny, které mají menší a větší hodnotu znaku

POZOR: soubor je třeba vždy uspořádat

pořadí prvku (kolikátý prvek to je, hodnota prvku je medián!)

určují vzorce :

Příklad

## **Kvantily**

Medián je kvantil dělící soubor na dvě poloviny dle předch. pravidel  
obdobně

kvantily – na čtvrtiny,  $x_{25}$ ,  $x_{50}$ ,  $x_{75}$ ,

decily

percentily

kvantily obecně široké použití ve statistice a v geografii

## **Geografický medián**

Geografický medián je čára dělící plochu, kde se jev vyskytuje tak, aby hodnota jevu byla v obou plochách stejná

## **Charakteristiky variability**

### **Variační rozpětí**

rozdíl největší a nejmenší hodnoty sledovaného statist. znaku

$$R = x_{\max} - x_{\min}$$

jednoduchá charakteristika

podléhá extrémním hodnotám, které mohou být i chybami

### **Průměrné odchyly**

průměrná odchyly je definována jako aritmetický průměr jednotlivých hodnot znaku od vybrané střední hodnoty (tj. od aritmetického průměru, mediánu, modu apod.)

### **Kvantilové odchyly**

Založeny na kladných odchylykách jednotlivých sousedních kvantilů

např. kvartilová odchylyka

decilová odchylyka

percentilová odchylyka

### **Střední diference**

je def. jako aritmetický průměr absolutních hodnot všech možných rozdílů jednotlivých hodnot sledovaného znaku

v praxi vhodná pouze pro malé soubory

### **Rozptyl a směrodatná odchylyka**

nejdůležitější charakteristiky variability

**Rozptyl  $s^2$**  z  $n$  hodnot znaku  $x$  je průměr druhých mocnin odchylek jednotlivých hodnot znaku od aritmetického průměru

**směrodatná odchylyka  $s$**  je mírou měnlivosti hodnot souboru kolem aritmetického průměru je druhou odmocnina rozptylu

### **Variační koeficient**

je častou používanou relativní mírou variability

je definován jako poměr směrodatné odchylyky k aritmetickému průměru

## **Charakteristiky asymetrie**

- Charakteristiky asymetrie ( míry šikmosti) jsou čísla dávající představu
- **souměrnosti tvaru** rozdělení četností
- míra šikmosti pro souměrné rozdělení je nula
- pro nesouměrné je kladná nebo záporná

## **Charakteristiky špičatosti( míry špičatosti)**

jsou čísla charakterizující koncentraci prvků souboru v blízkosti určité hodnoty znaku