

GEOSTATISTIKA - cv. 6: Statistický popis prostorového uspořádání bodů

Zadání:

Zjistěte, jestli je drogová kriminalita v londýnských čtvrtích Camden Town (with Primrose Hill), West End a Harlesden prostorově náhodně rozložená nebo se váže na určitou oblast. Pomocí vhodné charakteristiky popište, k jakému z teoretických rozložení (shlukové či pravidelné) se vámi zjištěné uspořádání ve čtvrtích blíží (udejte statistickou významnost). Stručně interpretujte hodnoty vypočtených charakteristik. K hodnocení prostorového uspořádání sídel použijte **metodu nejbližšího souseda a kvadrátovou analýzu**.

Vstupní data:

Pro práci použijte datasety, které najdete v ISu:

- **London_LAU2.shp** – administrativní dělení Velké Británie v úrovni LAU2
- **London_Crimi_2016-02.shp** – bodová hlášení přestupků a trestných činů v únoru 2016 ("Crime_type" = 'Drugs').

Pracujte v souřadném systému ETRS 1989 UTM Zone 30N.

Poznámky:

- **Analýza nejbližšího souseda** je založena na porovnání pozorované průměrné vzdálenosti mezi nejbližšími sousedy (r_{obs}) a průměrné vzdálenosti u známého vzorku (pattern) – tedy očekávané (r_{exp}). Pozorovaná průměrná vzdálenost mezi nejbližšími sousedy může být větší či menší než vzdálenost při náhodném rozmístění bodů. Používaná statistika je poměrem výše uvedených vzdáleností:

$$R = \frac{r_{obs}}{r_{exp}}$$

Interpretace R-statistiky:

Čím je hodnota $R < 1$, tím více se prostorové rozložení bodů blíží rozložení shlukovému ($r_{obs} < r_{exp}$).

Čím je hodnota $R > 1$, tím více se prostorové rozložení bodů blíží rozložení pravidelnému ($r_{obs} > r_{exp}$).

- **K významu vypočtených parametrů:** Program poskytuje hodnoty vypočtené a očekávané nejbližší vzdálenosti, dále R-statistiku. Standardizovaná hodnota (Z_R z-score) slouží k testování statistické významnosti:

Je-li $Z_R < -1,96$ či $Z_R > 1,96$ potom vypočtený rozdíl mezi pozorovaným a náhodným uspořádáním je statisticky významný – tedy není náhodný a naopak.

Zadaný úkol vyřešte pro všechny 3 čtvrtě metodou nejbližšího souseda v prostředí ArcMap. Metodu lze spustit pomocí ArcToolbox – Spatial Statistics Tools – Analyzing Patterns – Average Nearest Neighbour.

- **Kvadrátová analýza** – nejprve je nadefinována síť kvadrátů (čtverců). Kvadráty generujte pomocí nástroje **Create Fishnet** přibližně v rozsahu čtvrtě. Tato síť se přeloží

přes studovanou oblast. Pro nadefinování sítě musíte určit počet buněk v síti. **První varianta:** vyzkoušejte výpočet, kdy počet buněk je roven přibližně polovině počtu bodů. **Druhá varianta:** vyzkoušejte postup, kdy velikost jedné buňky a počet buněk jsou odvozeny z následujících vztahů: Optimální velikost kvadrátů (QS) lze získat ze vztahu:

$$QS = \frac{2 \cdot A}{n}$$

kde A je plocha studované oblasti a n počet analyzovaných bodů.

Velikost strany jedné buňky je potom

$$\sqrt{2A/n}$$

- Zjistěte, kolik sídel je v každém kvadrantu (**Spatial Join**).
- Pro rozdělení uvažujte tyto modelové hodnoty (P značí pravidelné rozdělení):

Počet sídel ve čtverci	P Camden Town	P West End	P Harlesden	Shlukové
0	16	11	13	39
1	20	17	15	0
2	7	3	6	0
3	3	2	2	0
4	1	1	1	0
...
SUMA	0	0	0	1

- Pomocí materiálů z přednášky vypočtete testovací kritéria a kritickou hodnotu.
- Interpretace: Stejně jako v obecném postupu testování porovnáváte vypočtené a kritické hodnoty testovacího kritéria. Je-li vypočtená hodnota vyšší než kritická, potom se dané uspořádání bodů statisticky významně liší (na zvolené hladině) od uspořádání náhodného.