

GEOSTATISTIKA - cvičení č. 3: Statistický popis prostorového uspořádání bodů

Zadání:

Charakterizujte prostorové uspořádání 20-ti sídel s nejvyšším počtem obyvatel ve vámi zvoleném okrese. Otestujte, zda existuje statisticky významný rozdíl mezi vámi zjištěným uspořádáním a uspořádáním náhodným.

Pomocí vhodné charakteristiky popište, k jakému z teoretických rozložení (shlukové či pravidelné) se vámi zjištěné uspořádání blíží (udejte statistickou významnost). Stručně interpretujte hodnoty vypočtených charakteristik.

K hodnocení prostorového uspořádání sídel použijte těchto metod

- Kvadrátová analýza (ArcView 3.2)
- Analýza nejbližšího souseda (ArcView 3.2 nebo ArcMap)

Poznámky:

- Využijte datových souborů vytvořených v cvičení 1.
- K vypracování v prostředí ArcView využijte projektu Ch3.apr, který naleznete na disku **E: (DATA)** ve složce **Geostatistika\Cviceni_3**
- Výše uvedené prostorové statistiky naleznete v nabídce **Point patterns**.
- **Quadrat analysis** – nejprve je nadefinována síť kvadrátů (čtverců). 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

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

- Dále vypočtené statistiky mají následující význam: **Length** – délka strany buňky, dále počet řádků a sloupců sítě. **Lambda** – průměrný počet bodů ve čtverci. Následně **četnosti** bodů v buňkách (frequency), hodnota **rozptylu** (variance), hodnota **poměru** (rozptyl/průměr), **testovací kritérium K-S testu**. Poté následuje volba hladiny významnosti (volte 0,05) a nakonec **kritická hodnota K-S testu** pro zvolenou hladinu významnosti.
- 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.

- **Ordered Neighbour Statistics** – analýza nejbližšího souseda. Metoda 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:

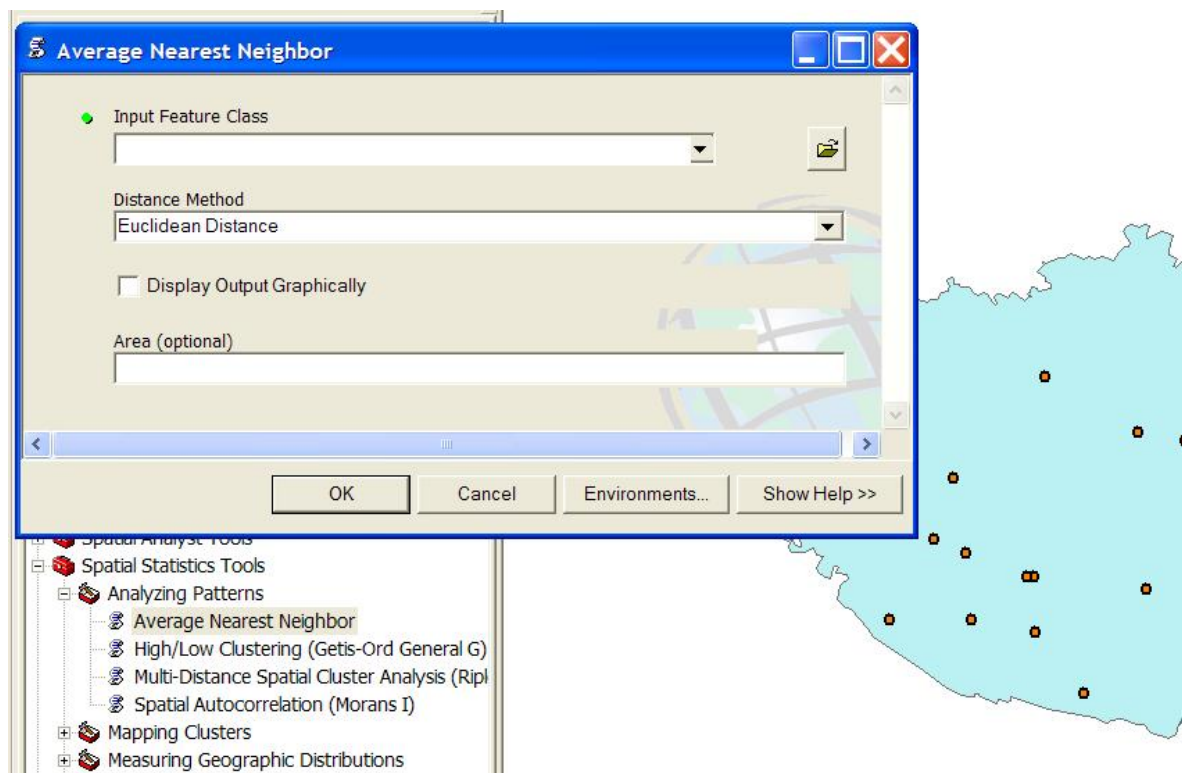
Čím je hodnota $R < 1$, tím více se prostorové rozložení bodů blíží rozložení shlukovému ($r_{obs} < r_{exp}$).

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 metodou nejbližšího souseda také v prostředí ArcMap. Metodu lze spustit pomocí **ArcToolbox – Spatial Statistics Tools – Analyzing Patterns – Average Nearest Neighbour** (viz. obr).



V jakých situacích by bylo vhodné využít různých metod výpočtu vzdálenosti (Distance Method)? Jakým způsobem ovlivňuje výsledky nastavení parametrů prostředí (Environment settings)?

Výsledkem cvičení je mapa obsahující vstupní datové vrstvy, vypočtenou prostorovou statistiku, veškeré náležitosti mapové kompozice (název, měřítko, legenda, ...) a dále slovní vyhodnocení úkolu.