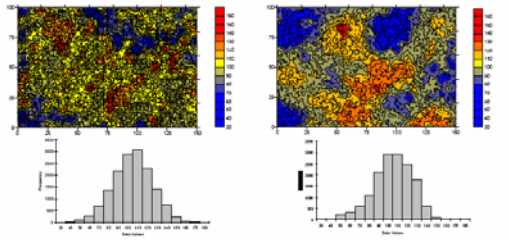
GEOSTATISTIKA – vymezení pojmu

• V užším slova smyslu – skupina interpolačních algoritmů založených na metodě krigingu.
• V širším slova smyslu – statistická analýza prostorově lokalizovaných dat.

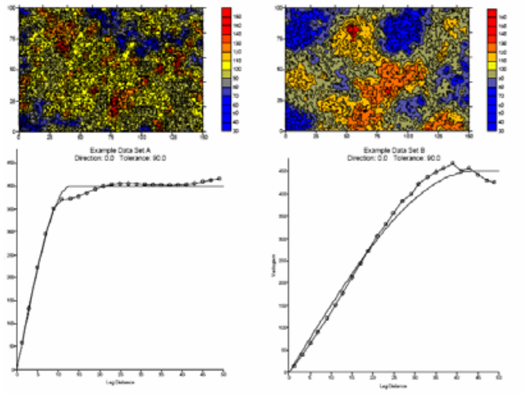
Pomocí „klasických“ statistických metod lze vhodně analyzovat především atributová data – jejich kvantitativní či kvalitativní vlastnosti. Velmi omezeně však jimi lze charakterizovat prostorové vlastnosti objektů a jevů.

Prostorové vlastnosti jako např. spojitost jevů, prostorovou autokorelaci, prostorové uspořádání (strukturu) lze charakterizovat právě pomocí geostatistických metod.



	A	B
Count	15251	15251
Average	100.00	100.00
Standard Deviation	20.00	20.00
Median	100.35	100.92
10 Percentile	73.89	73.95
90 Percentile	125.61	124.72

Prezentace prostorového rozšíření spojitého jevu metodami popisné statistiky



Prezentace prostorového rozšíření spojitého jevu metodami geostatistiky (tzv. semivariogram)

GEOSTATISTIKA – vymezení pojmu

- Statistický popis prostorově lokalizovaných dat (geografických objektů)
- Statistický popis prostorového uspořádání objektů (bodů, linií, ploch)
- Koncept prostorové autokorelace
- Strukturní analýzu a popis prostorové autokorelace strukturními funkcemi
- Konstrukci spojitých polí metodami krigingu
- Objektivní metody klasifikace jevů

Statistický popis bodů

Body představují nejčastější způsob prezentace geografických jevů. Body jsou zpravidla umísťovány v těžišti objektů.

To, jaké geografické objekty lze popsat pomocí bodů (tedy stupeň abstrakce) závisí na měřítku, ale také na druhu analýzy

Např. pro modelování optimálního spojení v síti sídel je vhodné je prezentovat centroidem, který tvoří uzel sítě).

Výpočet popisné statistiky často předchází použití geostatistických metod.

Umožňuje totiž ověřit některé vlastnosti studovaných souborů, které jsou pro aplikaci metod geostatistiky nezbytné. Jedná se o ověření takových vlastností jako je normalita rozdělení, stacionarita, linearita vztahu dvou veličin apod.

Charakteristiky polohy

Charakteristiky polohy slouží k určování geografického středu či mediánu.

Průměrný střed (mean centre)

Průměrný střed leží na průměru souřadnic X a Y. Má stejné nevýhody jako aritmetický průměr – je to především citlivost na extrémní hodnoty. Například v případě shlukového uspořádání bodů průměrný střed dobře nereprezentuje množinu bodů.

$$(\bar{x}_{mc}, \bar{y}_{mc}) = \left(\frac{\sum_{i=1}^{n} x_i}{n}, \frac{\sum_{i=1}^{n} y_i}{n} \right)$$

Kde x_i, y_i jsou souřadnice bodu i a n je počet bodů.

Vážený průměrný střed (weighted mean centre)

Používá se v případě výskytu více událostí/objektů na stejném místě. Pak má každý bod váhu přímo úměrnou počtu událostí/objektů na tomto místě.

Např. při výpočtu prostorového průměru několika měst bude průměrný střed dávat realističtější představu o centrální tendenci jestliže ho budeme vážit počtem obyvatel jednotlivých měst (nebo – koncentrací znečišťující látky v jednotlivých místech či frekvencí výskytu určitého jevu).

$$(\bar{x}_{wmc}, \bar{y}_{wmc}) = \left(\frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}, \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i} \right)$$

kde w_i jsou váhy jednotlivých bodů.

Mediánový střed (Median Center)

1. najdeme medián na ose X a Y a vedeme z nich linie kolmé na směr osy. Takto definovaný „medián ze souřadnic“ ale nemusí odpovídat mediánu souboru bodů, protože distribuce nemusí být mezi kvadranty vyrovnaná.
2. (UK) - Mediánový střed je střed, kterým se studovaná plocha dělí do čtyř kvadrantů, z nichž každý obsahuje stejný počet bodů.
3. (US) - Mediánový střed jako střed vyžadující minimální (nejkratší) cestu. Tj. celková vzdálenost z mediánového středu do každého z bodů je minimální. Jinak řečeno – cesta z jakéhokoliv jiného místa do všech bodů oblasti bude delší než cesta z mediánového středu. Tuto podmínku lze vyjádřit vztahem:

$$\min \sum \sqrt{(x_i - u)^2 + (y_i - v)^2}$$

kde x_i a y_i jsou souřadnice jednotlivých bodů a u, v jsou souřadnice mediánového středu. Analogickým způsobem lze definovat tzv. vážený mediánový střed:

$$\min \sum f_i \sqrt{(x_i - u)^2 + (y_i - v)^2}$$

Váhy f_i pro jednotlivé body mohou být negativní či pozitivní podle toho, zda daný bod přitahuje či naopak odpuzuje polohu mediánového středu

Charakteristiky rozptylu

Směrodatná vzdálenost (*standard distance*)

$$SD = \sqrt{\frac{\sum_{i=1}^{n} (x_i - x_{mc})^2 + \sum_{i=1}^{n} (y_i - y_{mc})^2}{n}}$$

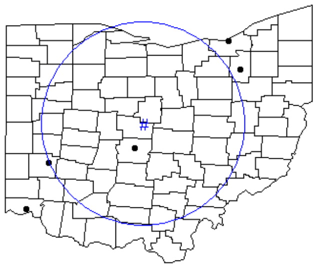
Vážená směrodatná vzdálenost (weighted standard distance)

$$SD = \sqrt{\frac{\sum_{i=1}^{n} f_i (x_i - x_{mc})^2 + \sum_{i=1}^{n} f_i (y_i - y_{mc})^2}{\sum_{i=1}^{n} f_i}}$$

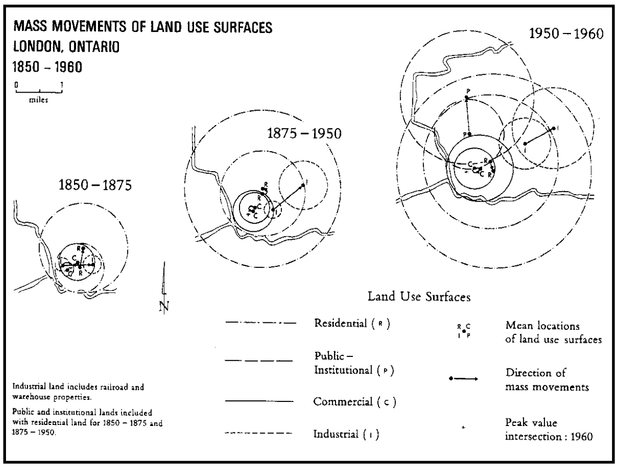
Směrodatná vzdálenost je nejčastěji používána ve formě kružnice kolem průměrného středu (**Standard distance circle**), jejíž poloměr je právě hodnota směrodatné vzdálenosti.

Různé směrodatné vzdálenosti pro různý typ jevů lze zakreslovat do stejného území. Tyto kružnice nám dávají představu o rozptylu hodnot kolem střední hodnoty pro jednotlivé typy jevů.

Mohou být použity i pro studium dynamiky jevů (různé kružnice pro jeden jev v různých časových horizontech).



Poloha váženého průměrného středu a kružnice směrodatné vzdálenosti pro pět měst ve státě Ohio. Jako váhy byl použit počet obyvatelstva



Směrodatná vzdálenost největších měst vážená počtem jejich obyvatel

SD = 8,849 **SD = 3,277**

Absolutní standardní vzdálenost lze vážit plochou porovnávaných států:

0,027 **0,238**

Modelování prostorového uspořádání bodů

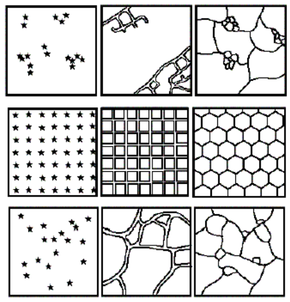
Rozmístění bodů v prostoru je výsledkem určitých procesů či vhodných podmínek (lokace měst je výsledkem působení faktorů jako reliéf, přírodní zdroje, komunikace, atd.)

Cílem studia prostorového rozmístění bodů je zjistit:

- **jak daleko má konkrétní rozmístění objektů k rozmístění teoretickému**
- **jak se liší rozmístění bodů ve dvou různých oblastech**
- **jak se mění rozmístění bodů v rámci jedné oblasti v čase.**

Statisticky prokázaný výskyt určitého prostorového uspořádání může být základem pro zjišťování příčin, které vedly k pozorovanému uspořádání.

Základní typy prostorového uspořádání bodů



- **Shlukové (Clustered)**
- **Pravidelné (Regular)**
- **Náhodné (Random)**

Základní metody statistického popisu prostorového uspořádání bodů

- **Analýza kvadrátů** – testujeme, zda rozmístění bodů v ploše je náhodné či nikoliv.
- **Metoda nejbližšího souseda** – porovnává průměrnou vzdálenost mezi nejbližšími sousedy pole bodů k teoretickému rozmístění.
- **Prostorová autokorelace** – měří jak podobné či nepodobné jsou hodnoty atributů sousedních bodů.

Problémy spojené s popisem prostorového uspořádání bodů

- měřítko
- rozsah studované oblasti
- kartografická projekce

Měřítko – je nutné vhodně zvolit tak, aby studovaný jev mohl být prezentován body v prostoru.

Rozsah studované oblasti

V závislosti na zvolené oblasti (často vymezené administrativními hranicemi) se mění jak vzdálenosti mezi jednotlivými body, tak také charakteristiky jejich prostorového uspořádání.

Kartografická projekce

Projekce se volí podle účelu (viz. analýza kvadrátů).

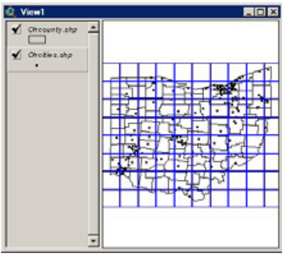
Projekcí se mění tvar, vzdálenosti, vzájemná poloha objektů.

Čím větší studovaná oblast, tím větší bude role zvolené projekce.

Analýza kvadrátů (QUADRAT ANALYSIS)

• Je založena na **hodnocení změn hustoty bodů** v prostoru. Je porovnáváno, zda rozmístění bodů v prostoru je náhodné, či má blíže k uspořádání shlukovému či pravidelnému.

• Studovaná plocha je rozdělena pravidelnou sítí na buňky a je zjištěn počet bodů v každé buňce.



Analýza kvadrátů

• Je analyzováno rozdělení četností buněk s určitým počtem bodů.

• Toto rozdělení je porovnáváno s náhodným rozdělením četností.

• **Extrémně shlukové uspořádání** – většina bodů v jedné či několika málo buňkách.

• **Extrémně pravidelné** – ve všech buňkách přibližně stejně

• Buňky se označují jako **kvadráty** a nemusí jít o čtverce, ale např. i o kruhy či šestiúhelníky – je to dáno empirií.

• V rámci jedné analýzy však tvar a velikost buněk musí být konstantní.

Optimální velikost kvadrátů (QS)

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

A - plocha studované oblasti

n - počet analyzovaných bodů.

Velikost strany vhodného kvadrátu

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

Získané rozložení četností bodů v kvadrátech (empirické) je porovnáváno s náhodným rozložením (teoretickým). **Vhodným testem je např. K-S test**.

Testování výsledků analýzy kvadrátů

Počet měst v každém čtverci	Zjištěné rozdělení	Pravidelné rozdělení	Shlukové rozdělení
0	36	0	79
1	17	26	0
2	10	26	0
3	3	26	0
4	2	2	0
5	2	0	0
6	1	0	0
7	1	0	0
8	1	0	0
9	1	0	0
10	1	0	0
11	1	0	0
12	1	0	0
13	1	0	0
14	1	0	0
28	1	0	0
164	0	0	1

Zjištěné rozdělení četnosti 164 měst v kvadrátech ve studované oblasti a rozdělení četností teoretická

Praktický postup testování výsledků analýzy kvadrátů

1. (H0) - neexistuje statistiky významný rozdíl (je-li rozdíl malý, může být výsledkem náhody, čím je větší, s tím větší pravděpodobností náhodný není, ale je statistiky významný).

2. Zvolíme hladinu významnosti α = 0,05

3. Vypočteme kumulované četnosti

4. Vypočteme testovací kritérium: $D = \max|O_i - E_i|$

5. Vypočteme kritickou hodnotu: $D_\alpha = \frac{1{,}36}{\sqrt{m}}$ $\quad D_\alpha = 1{,}36\sqrt{\frac{m_1 + m_2}{m_1 \cdot m_2}}$

6. Je-li vypočtená hodnota D větší než kritická hodnota D_α, potom rozdíl mezi oběma uspořádáními je statisticky významný.

Testování výsledků analýzy kvadrátů K-S testem

Počet měst v každém čtverci	Zjištěné rozdělení	Relativní četnosti	Kumulativní četnosti	Pravidelné rozdělení	Relativní četnosti	Kumulativní četnosti	Absolutní diference
0	36	0,450	0,450	0	0,000	0,00	0,45
1	17	0,213	0,663	26	0,325	0,33	0,34
2	10	0,125	0,788	26	0,325	0,65	0,14
3	3	0,038	0,825	26	0,325	0,98	0,15
4	2	0,025	0,850	2	0,025	1,00	0,15
5	2	0,025	0,875	0	0,000	1,00	0,13
6	1	0,013	0,888	0	0,000	1,00	0,11
7	1	0,013	0,900	0	0,000	1,00	0,10
8	1	0,013	0,913	0	0,000	1,00	0,09
9	1	0,013	0,925	0	0,000	1,00	0,08
10	1	0,013	0,938	0	0,000	1,00	0,06
11	1	0,013	0,950	0	0,000	1,00	0,05
12	1	0,013	0,963	0	0,000	1,00	0,04
13	1	0,013	0,975	0	0,000	1,00	0,03
14	1	0,013	0,988	0	0,000	1,00	0,01
28	1	0,013	1,000	0	0,000	1,00	0,00
164	0	0,000	1,000	0	0,000	1,00	0,00

Testovací kritérium: D = 0,45

Kritická hodnota pro α = 0,05: D_α = 0,2115

Zamítáme nulovou hypotézu - rozdělení měst se statisticky významně liší od rozdělění pravidelného

Testování výsledků analýzy kvadrátů

Poissonovo rozdělení (Poisson random process) je určeno průměrnou frekvencí výskytu (λ) v jednotlivých jednotkách (kvadrátech):

$$\lambda = {n}/{m}$$

m – počet kvadrátů; *n* - počet bodů v prostoru

Hodnoty průměru a rozptylu Poissonova rozdělení se rovnají hodnotě (λ).

Bude-li distribuce bodů v prostoru generována náhodným procesem, potom toto rozdělení má stejný průměr a rozptyl a jejich poměr se bude blížit jedné.

Postup testování

1. Vypočteme hodnoty průměru a rozptylu pro četnosti bodů v kvadrátech.
2. Hodnoty dáme do poměru, hodnotu porovnáme s 1.
3. Rozdíl lze standardizovat (vyjádřit v násobcích směrodatné odchylky).
4. Vyjde-li hodnota větší než 1,96, potom je rozdíl statisticky významný na hladině α = 0,05.

Test založený na poměru průměru a rozptylu je silnější než K-S test

Lze ho však použít pouze v případě, že předpokládáme Poissonovo rozdělení studované množiny bodů.

Testování výsledků analýzy kvadrátů vůči rozložení generovanému Poissonovým náhodným procesem (pro λ = 2,05)

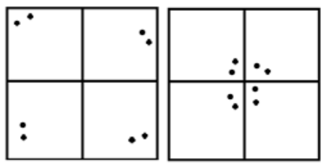
Počet měst v každém čtverci	Zjištěné rozdělen	Relativní četnosti	Kumulativní četnosti	Pravděpodobnosti Poissonova rozd.	Kumulativní četnosti	Absolutní diference
0	36	0,450	0,450	0,1287	0,13	0,32
1	17	0,213	0,663	0,2639	0,39	0,27
2	10	0,125	0,788	0,2705	0,66	0,12
3	3	0,038	0,825	0,1848	0,85	0,02
4	2	0,025	0,850	0,0947	0,94	0,09
5	2	0,025	0,875	0,0388	0,98	0,11
6	1	0,013	0,888	0,0133	0,99	0,11
7	1	0,013	0,900	0,0039	1,00	0,10
8	1	0,013	0,913	0,001	1,00	0,09
9	1	0,013	0,925	0,0002	1,00	0,07
10	1	0,013	0,938	0	1,00	0,06
11	1	0,013	0,950	0	1,00	0,05
12	1	0,013	0,963	0	1,00	0,04
13	1	0,013	0,975	0	1,00	0,02
14	1	0,013	0,988	0	1,00	0,01
28	1	0,013	1,000	0	1,00	0,00
164	0	0,000	1,000	0	1,00	0,00

Testovací kritérium: D = 0,3213

Kritická hodnota pro α = 0,05: D_{α} = 0,1520

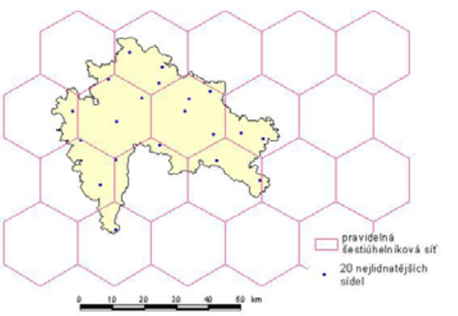
Rozdíl mezi oběma uspořádáními je statisticky významný.

Omezení analýzy kvadrátů:



Analýza kvadrátů neřeší otázku rozložení bodů uvnitř kvadrátů.

Prostorové uspořádání nejlidnatějších sídel okresu Jindřichův Hradec
Pravidelná šestiúhelníková síť použitá pro kvadrátovou analýzu



Délka strany buňky (šestiúhelník) = 11433,4 m
Lambda (průměrný počet bodů v šestiúhelníku) = 1, Rozptyl = 1,6
Poměr (rozptyl/průměr) = 1,84932
Testovací kritérium K-S testu = 0,182121
Kritická hodnota K-S testu (na hladině významnosti 0,05) = 0,294

Analýza nejbližšího souseda

(NEAREST NEIGHBOUR ANALYSIS)

Metoda analýzy kvadrátů je založena na konceptu **hustoty** (počet bodů v ploše)

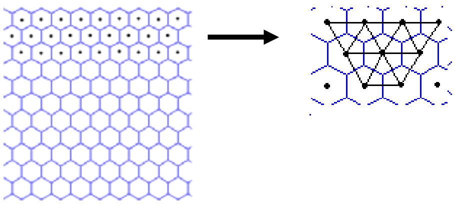
Metoda analýzy nejbližšího souseda je naopak založena na konceptu **vzdálenosti** (spacing – plocha připadající na bod).

Metoda analýzy nejbližšího souseda je založena na porovnání pozorované průměrné vzdálenosti mezi nejbližšími sousedy a této průměrné vzdálenosti u známého (teoretického) prostorového uspořádání (pravidelného či náhodného).

Pravidelné uspořádání bodů

Nejpravidelnější uspořádání – studovaná oblast je rozdělena sítí pravidelných šestiúhelníků a body v této oblasti tvoří jejich středy.

Body lze pospojovat do sítě pravidelných trojúhelníků.



R statistika (R – randomness)

Určí se jako poměr mezi pozorovanou a očekávanou průměrnou vzdáleností nejbližších sousedů v určité oblasti:

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

Hodnotu r_{obs} zjistíme tak, že určíme vzdálenost mezi daným bodem a všemi jeho sousedy. Dále najdeme nejkratší vzdálenost – tedy nejbližšího souseda. Tento proces se opakuje pro všechny body. Ze všech nejkratších vzdáleností se vypočte průměr.

Hodnotu r_{exp} zjistíme ze vztahu:

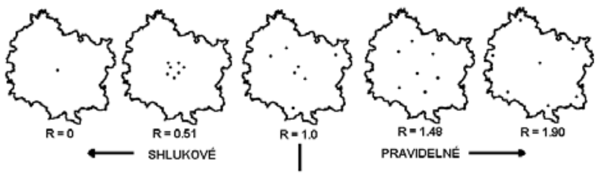
$$r_{\exp} = \frac{1}{2\sqrt{n/A}}$$

kde *A* je plocha a *n* počet bodů v ploše.

Interpretace hodnot 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 ($r_{obs} > r_{exp}$).



R = 0	zcela shlukové uspořádání
R = 1	náhodné uspořádání
R = 2,149	zcela pravidelné uspořádání

K hodnocení rozdílu mezi pozorovanou a očekávanou vzdáleností nejbližšího souseda lze využít tzv. **směrodatné chyby** (Standard Error – SE_r)

$$SE_r = \frac{0{,}26136}{\sqrt{n^2/A}}$$

Směrodatná chyba popisuje pravděpodobnost, že jakýkoliv rozdíl dvou hodnot je výsledkem náhodných vlivů. Je-li zjištěná diference malá ve srovnání s *SE*, potom rozdíl není statisticky významný a naopak.

Za statisticky významný považujeme rozdíl, který můžeme obdržet v pěti případech ze sta – tedy s pravděpodobností 5 %, α=0,05.

Vyjádřeno v násobcích směrodatné chyby - rozdíl mezi dvěma populacemi považujeme za statisticky významný, jestliže je menší než -$1{,}96SE_r$ a nebo větší než $+1{,}96SE_r$:

Pravděpodobnost (<95%) = (-1,96SE_r; +1,96SE_r)

Standardizace hodnot rozdílů

Pomocí směrodatné chyby lze vypočítat standardizovanou hodnotu (Z-score):

$$Z_R = \frac{r_{obs} - r_{\exp}}{SE_r}$$

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

Problémy spojené s metodou analýzy nejbližšího souseda:

- Nelze spoléhat na vizuální srovnání prostorového rozložení ani na vypočtenou hodnotu R. Ta by měla být doplněna hodnotou Z_R pro ověření statistické významnosti pozorovaného rozdílu.
- Metoda analýzy nejbližšího souseda může být rozšířena na analýzu nejbližších sousedů druhého, třetího a vyšších řádů.
- Výsledky jsou vysoce citlivé k měřítku (lokální vs. regionální)
- V závislosti na studovaném jevu musí být věnována pozornost vymezení studované plochy (administrativní či přirozené hranice).

Modelování prostorového uspořádání bodů s využitím prostorové autokorelace

(SPATIAL AUTOCORRELATION)

Koncept prostorové autokorelace

Jak analýza kvadrátů tak analýza vzdálenosti nejbližšího souseda pracují pouze s polohou bodů. Nerozlišují body podle hodnot jejich atributů.

Oba parametry (polohu i atributy) hodnotí **prostorová autokorelace** (SA) – je tedy metodou vhodnější.

Východiska prostorové autokorelace: Většina jevů se v prostoru mění spojitě. Blízké body budou mít i podobné hodnoty studovaného jevu a naopak.

(First law of geography - Tobler, 1970)

Koeficienty prostorové autokorelace

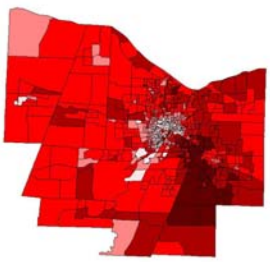
Míry prostorové autokorelace kombinují v jednom výrazu **míry podobnosti atributů** i **míry podobnosti polohy.**

Mezi nejpoužívanější koeficienty prostorové autokorelace náleží:

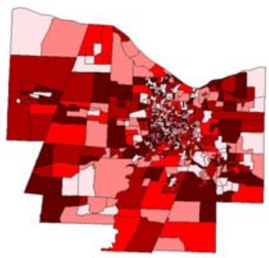
- **Gearyho poměr C (Geary's Ratio)**
- **Moranův index I (Moran's I)**

Lze jich využít pro intervalová a poměrová data.

Moranův index I - příklad



Průměrný příjem
Moran's I: 0,66



Náhodná proměnná
Moran's I: 0,012

Míry prostorové autokorelace

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} w_{ij}$$

c_{ij} – podobnost atributu v bodě *i* a *j*

w_{ij} – vzdálenost bodu *i* a *j*. $w_{ii} = 0$ pro všechny body

x_i – hodnota studovaného atributu v bodě *i*

n – počet bodů ve vyšetřovaném vzorku

Koeficienty prostorové autokorelace

Koeficient prostorové autokorelace - SAC (spatial autocorrlelation coefficient) je úměrný vážené míře podobnosti atributů bodů – obecně:

$$SAC \approx \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} \cdot w_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}}$$

Gearyho poměr C:

V případě Gearyho poměru se podobnost hodnot atributu mezi dvěma body vypočte podle následujícího vztahu:

$$c_{ij} = (x_i - x_j)^2$$

$$C = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} \cdot w_{ij}}{2\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \cdot \sigma^2} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \cdot (x_i - x_j)^2}{2\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \cdot \sigma^2}$$

kde $\boldsymbol{\sigma^2}$ je rozptyl hodnot atributu *x* s průměrem $\overline{x}$

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{(n-1)}$$

Moranův index I

V případě **Moranova indexu** se podobnost hodnot atributu v bodech *i* a *j* vyjádří následovně: $c_{ij} = (x_i - \overline{x}) \cdot (x_j - \overline{x})$

$$I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} \cdot w_{ij}}{s^2 \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \cdot (x_i - \overline{x}) \cdot (x_j - \overline{x})}{s^2 \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}$$

kde s^2 je v tomto případě výběrový rozptyl:

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n}$$

Definování míry podobnosti polohy bodů

Podobnost polohy bodů *i* a *j*, - hodnota w_{ij}. se určí jako inverzní hodnota vzdálenosti těchto bodů.

Tedy podle výše uvedených předpokladů dáváme malou váhu hodně vzdáleným bodům a velkou váhu bodům blízkým tedy:

$$w_{ij} = 1/d_{ij}$$

Modifikovat lze také váhy vzdálenosti bodů výrazem:

$$w_{ij} = 1/d_{ij}^{b}$$

kde koeficient *b* může nabývat různých hodnot v závislosti na povaze studovaného problému (vzdálenost měřená dosažitelností autem a letadlem je jiná). Hodnota *b* je často rovna 2.

Obor hodnot koeficientů prostorové autokorelace

Rozdíly mezi oběma indexy jsou dány způsobem výpočtu rozdílů mezi hodnotami atributu. Obor hodnot, kterých mohu oba indexy nabývat se tedy také liší, jak uvádí následující tabulka:

Prostorové uspořádání	Gearyho poměr C	Moranův index I
Shlukové uspořádání, sousední body vykazují podobné hodnoty	0 < C <1	I >E(I)
Náhodné uspořádání, body nevykazují znaky podobnosti	C ~ 1	I E(I)
Pravidelné uspořádání, sousední body vykazují rozdílné charakteristiky	1 < C < 2	I < E(I)

kde E(I) = (-1)/(n-1) je očekávaná hodnota indexu

Praktický postup výpočtu indexů a jejich interpretce

1. Vypočteme empirické hodnoty indexů I a C
2. Zjistíme očekávané hodnoty indexů I a C
3. Vypočteme hodnoty rozptylů indexů I a C
4. Tyto hodnoty potřebujeme pro testování, zda se vypočtené hodnoty indexů C a I statisticky významně liší od náhodného uspořádání.
5. Vyjádříme tzv. standardizované hodnoty (Z-skore)
6. Porovnáme standardizované hodnoty s hodnotami kritickými a zhodnotíme statistickou významnost hodnot indexů I a C

Odhad očekávaných hodnot Moranova indexu I a hodnot rozptylu pro náhodné uspořádání

$$E_N(I) = E_R(I) = \frac{-1}{n-1}$$

$$VAR_N(I) = \frac{\left(n^2 S_1 - nS_2 + 3W^2\right)}{W^2(n^2-1)} - \left[E_N(I)\right]^2$$

$$VAR_R(I) = \frac{n\left[\left(n^2 - 3n + 3\right)S_1 - nS_2 + 3W^2\right]}{(n-1)(n-2)(n-3)W^2} - \frac{k\left[(n^2 - n)S_1 - nS_2 + 3W^2\right]}{(n-1)(n-2)(n-3)W^2} - \left[E_R(I)\right]^2$$

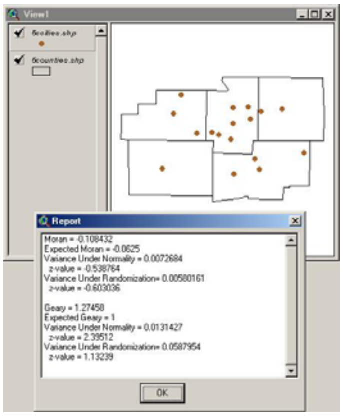
Určení standardizovaných hodnot

Máme-li vypočteny očekávané hodnoty indexů a jejich rozptyly, můžeme vyjádřit standardizované hodnoty (Z-skore)

$$Z = \frac{I - E(I)}{VAR(I)}$$

Pro hodnoty Z pak mohou být použity stejné kritické hodnoty, tedy na hladině významnosti α=0,05:

-1,96 < Z < +1,96

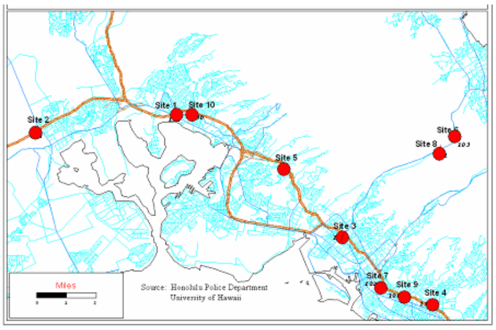


Příklad výpočtu měr prostorové autokorelace

Interpretace hodnot koeficientů prostorové autokorelace:

Pokud zjištěné hodnoty z-skóre padnou vně intervalu (-1,96 ; +1,96), potom se prostorové uspořádání bodů statisticky významně liší (na hladině 5 %) od uspořádání náhodného.

Příklady použití měr prostorové autokorelace



10 Highest Crash Frequency Intersections in Honolulu

(analýza lokálních a globálních vlivů)

Crime analysis (prostorová analýza kriminality)

http://www.ncjrs.gov/html/nij/mapping/index.html

http://www.icpsr.umich.edu/CRIMESTAT/

Figure 4.20

A map showing robbery hot spots, using STAC. West Bronx, New York.

Source: Block and Block, 1997, figure 7. Reproduced by permission.

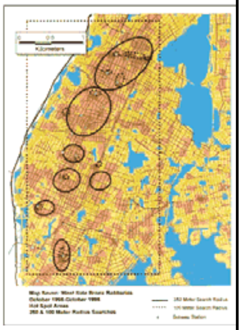


Figure 4.17

A map showing domestic burglary hot spots and police perception in the Meadows area of Nottingham, United Kingdom.

Source: Ratcliffe and McCullagh, 1998, figure 1, p. 48. Reproduced by permission.