Local Spatial Autocorrelation
Clusters

Luc Anselin

http://spatial.uchicago.edu
LISA principle

local Moran

local G statistics

issues and interpretation
LISA Principle
• Clustering vs Clusters

  global spatial autocorrelation does NOT suggest
  the location of the clusters

  cluster detection

  identification of location

  assessment of significance

  many cluster detection methods
• Local Indicators of Spatial Association

LISA (Anselin 1995)

local spatial statistic - one for each location

sum of LISA proportional to a corresponding global statistic
• Local Spatial Autocorrelation Analysis

assess significance of local statistic at each location

identification of location of spatial clusters (hot spots, cold spots) and spatial outliers

in absence of global S.A., or in presence of global S.A. (significance levels affected)
• LISA Forms of Global Statistics

• every decomposable statistic

if global = a. [ $\sum_i \text{component}(i)$ ]

then local = component(i)
Local Moran
Local Form of Moran’s I (Anselin 1995)

for row-standardized weights
(such that $S_0$ and $N$ cancel out in Moran’s I)

variables as deviations from mean ($z_i$)

$$ I_i = \frac{z_i}{m_2} \sum_j w_{ij} z_j $$

$$ m_2 = \sum_i z_i^2 $$ does not vary with $i$, thus constant

$$ I_i = \frac{1}{m_2} z_i \sum_j w_{ij} z_j = c. z_i \sum_j w_{ij} z_j $$
• Link Local-Global

\[ \sum_i I_i = N \cdot I \]

or: \[ I = \frac{\sum_i I_i}{N} \]

Global Moran is average of local Moran statistics
• Inference

analytical or computational

analytical approximation is poor (do not use)

computational based on conditional permutation
Conditional Permutation

conditional upon value observed at i

hold value at i fixed, random permute remaining n-1 values and recompute local Moran

repeat many times to obtain reference distribution

conditional permutation for each location
• **Local Significance Map**

shows locations with significant local statistic by level of significance

not very useful for substantive interpretation

diagnostic for sensitivity of results (for example, when only significant at 0.05)
local significance map
Nepal % not expected to survive past 40 (queen)
• Local Cluster Map

shows locations with significant local spatial autocorrelation by type of association

four color scheme

spatial clusters: high-high and low-low

spatial outliers: high-low and low-high

shown for a given level of significance (sensitivity analysis)
local cluster map for different p-values

- p < 0.05
- p < 0.01
high-low spatial outlier
• What is a Cluster?

- locations with significant positive local spatial autocorrelation are the core of a cluster

- actual “cluster” includes neighbors as well as core

- regions of high/low values rather than individual locations
local cluster map for $p < 0.001$ with neighbors highlighted
Local G Statistics
• **Local G Statistic**


not a LISA in a strict sense (no local-global connection) but useful for detecting clusters

based on point pattern analysis logic

two versions: \( G_i \) and \( G_i^* \) (value at \( i \) included)
$G_i$ Statistic

$$G_i = \Sigma_j w_{ij}x_j / \Sigma_j x_j \text{ for } j \text{ not equal } i$$

i not included in either numerator or denominator

numerator is weighted average of neighbors (spatial lag)

denominator is sum of all values, excluding the value of $x$ at i
Inference

analytical: based on an approximation

not very reliable

conditional permutation inference: same principle as for local Moran
Gi statistic significance map and cluster map
**G_i** Star Statistic

\[ G_i^* = \frac{\sum_j w_{ij} x_j}{\sum_j x_j} \text{ for all } j \]

i included in both numerator and denominator

numerator is weighted average of neighbors and value at i (need to define \( w_{ii} \))

denominator is sum of all values, thus constant

can be used as a local join count statistic
$G^* i$ statistic significance map and cluster map
• Interpretation

significant values only - ignore others

positive $G_i$ ($G_i^*$) = local clustering of high values hot spot

negative $G_i$ ($G_i^*$) = local clustering of low values cold spot

does NOT detect spatial outliers
Local Moran vs. G Statistics

G statistics useful when negative spatial autocorrelation is negligible (then hot spots and cold spots)

G statistics do not consider spatial outliers, local Moran does

Local Moran needs to be combined with classification of type of spatial autocorrelation
local Moran

local Gi*

local Moran’s I compared to local Gi*
Issues and Interpretation
• Multiple Comparisons

  significance level for a given location assumes only that location is being analyzed

  because all locations are analyzed, individual p-value is incorrect (too low)

  various corrections (e.g., Bonferroni bounds, false discovery rate) but none satisfactory

  in practice: cautious interpretation
• Exploratory Only

LISA clusters and outliers are identified, but not explained

suggests interesting locations

multiple processes can yield the same pattern
• **Univariate Only**

  univariate spatial autocorrelation can be due to other covariates

  univariate analysis ignores multivariate interactions

  scale mismatch can create impression of clusters without a meaningful process interpretation