Lesson 5: Advanced Spatial Point Patterns Analysis: Local Co-Location Quotient

Dr. Kam Tin Seong
Assoc. Professor of Information Systems(Practice)
School of Computing and Information Systems,
Singapore Management University

05 Feb 2023

Content

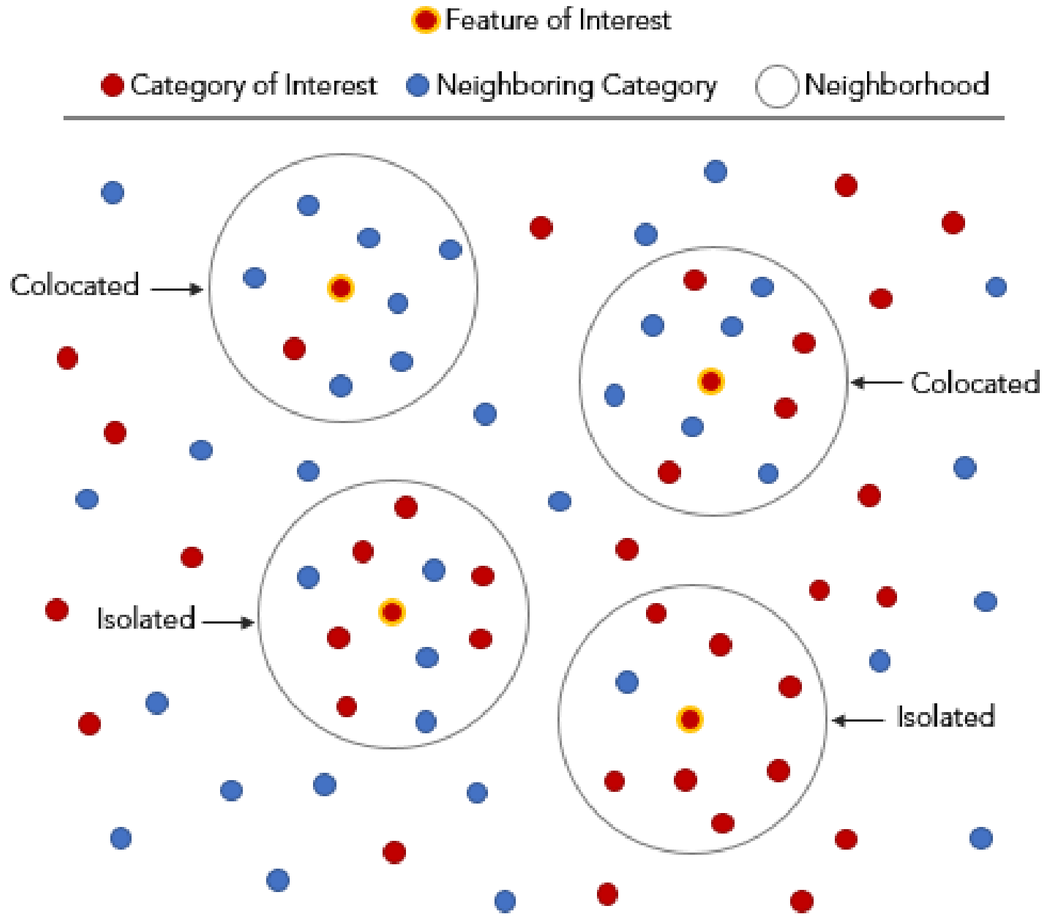
- Introduction to Local Co-Location Quotient (LCLQ)
- Basic Principles
- Application of LCLQ
- Mathematics of LCLQ
- Interpretation of LCLQ

Common Questions of Two Point Events

- Are certain business types likely to be colocated (such as coffee shops and retail stores)?
- Are locations of residential theft more likely to occur or be colocated with certain housing types?
- Are there specific areas in your study area where failed restaurant inspections are colocated with insect infestations?

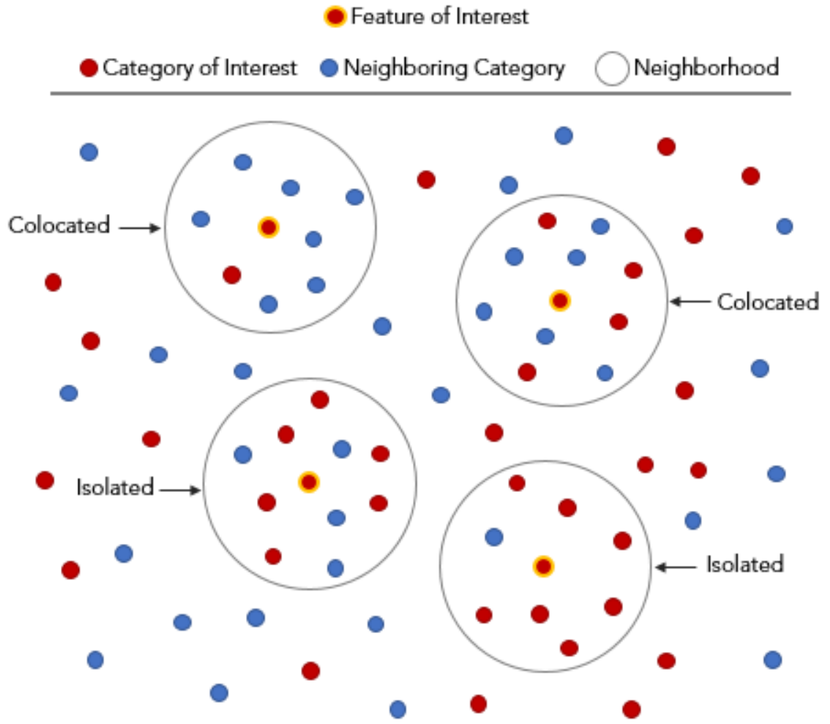
What is Colocated?

A point event category A is colocated with point events of category B if it is surrounded by several point event category B within a specified distance.



How the colocation quotient is calculated?

Each feature in the Category of Interest (category A) is evaluated individually for colocation with the presence of the Neighboring Category (category B) found within its neighborhood. In general, if the proportion of B points within the neighborhood of A is more than the global proportion of B, the colocation quotient will be high. If the neighborhood of A contains many other A points or many other categories other than B, the colocation between the Category of Interest (category A) and the Neighboring Category (category B) will be small.



The base formula

The local colocation quotient calculated from point Ai in the Category of Interest A to the Neighboring Category B is given as:

$$LCLQ_{Ai \to B} = \frac{N_{Ai \to B}}{N_B/(N-1)}$$

Where NB is the total number of category B present in the study area, and N is the total number of points in the study area (including all categories present). NAi–>B is the weighted average of the number of category B points in the neighborhood of each category A point (Ai). This is based on a distance decay function that allows closer features to the target feature to weigh heavier in the calculations than features that are farther away.

The wieghted average

NAi–>B represents the weighted average of the number of type B points in the neighborhood of each Ai based on either a Gaussian or Bisquare kernel function given as:

$$N_{A_i \to B} = \sum_{j=1(j \neq i)}^{N} \frac{w_{ij} * f_{ij}}{\sum_{j=1(j \neq i)}^{N} w_{ij}}$$

The kernel function

Where fij is a binary variable indicating whether point j is a category B point. If this is true, it is equal to 1. If not, it is equal to 0. The kernel function equations are given as:

$$For\ Gaussian\ Kernel, w_{ij} = \exp(-0.5 * \left(\frac{d_{ij}^2}{d_{ib}^2}\right))$$

$$For\ Bisquare\ Kernel^*, w_{ij} = 1 - \left(\frac{d_{ij}^2}{d_{ib}^2}\right)$$

> **Note**
>
> If the value of wij is negative for the Bisquare kernel, the weight assigned is 0.

Complete spatial randomness test

- Permutations are used to calculate a p-value for each of the Input Features of Interest to determine whether the observed colocation quotient values are statistically significant.
- For each of the features, the local colocation quotient is calculated using its neighborhood and for each permutation, the categories of all other points are randomly rearranged across the entire study area (keeping the target point location category constant).
- A new local colocation quotient is calculated for each feature of interest using the categories in the neighborhood for each permutation.
- The result is a reference distribution of colocation quotient values that is then compared to the actual colocation quotient value of the feature to determine the probability that the observed value could be found in the random distribution from the permutations.
- If the p-value is small (less than 0.05), the actual colocation quotient for the feature is statistically significant. The default for the tool is 99 permutations; however, the precision of the p-value calculated is improved with increased permutations.

Interpreting results

- When the Colocation Analysis tool is run, it adds two fields to the resulting Output Features. The **Local Colocation Quotient** field contains the resulting quotient score for each of the Input Features of Interest, and the **p-value** is also reported.
- Features of the Category of Interest (category A) that have a local colocation quotient greater than one are more likely to have features of the Neighboring Category (category B) within their neighborhood.
- Features that have colocation quotients less than one are less likely to have category B within their neighborhood.
- If a feature has a colocation quotient equal to one, it means the proportion of categories within their neighborhood is a good representation of the proportion of categories throughout the entire study area.