

複数種の点分布における空間的混合

東京大学 *貞広 幸雄 SADAHIRO Yukio

1. はじめに

同一地域に複数の点分布が存在する場合、点分布同士の空間的混合が発生することがある。混合の度合いは一様ではなく、場所によって様々である場合が多い。本論文では点分布間の空間的混合を定量的に評価し、混合パターンが他とは異なる地域を抽出する手法を提案する。

2. 手法

いま、領域 S に K 種類の点分布があり、 i 種の点の個数を N_i 、その割合を μ_i とする。位置 \mathbf{x} に中心を置く半径 r の円を $Z(\mathbf{x}, r)$ と書く。この円の内外で、各点が各種類に確率的に割り当てられるモデルを考える。円 $Z(\mathbf{x}, r)$ の内外で i 種になる確率をそれぞれ p_i 、 q_i とし、実際の点データから最尤法によってそれらの値を推定する。このとき、 $\mu_i > p_i$ であれば円 $Z(\mathbf{x}, r)$ 内で i 種の点が過少であり、 $\mu_i < p_i$ であれば過多であることになる。

次に、円の内外で点の混合度に差が無いというモデル、即ち、全ての i について $p_i = q_i$ というモデルを帰無仮説、推定されたモデルを対立仮説とし、対数尤度 $\lambda(Z(\mathbf{x}, r))$ によって各モデルの妥当性を評価する。ここで $\lambda(Z(\mathbf{x}, r))$ は、円 $Z(\mathbf{x}, r)$ 内における点分布の構成が、円の外側 $S - Z(\mathbf{x}, r)$ と比較してどの程度異なっているかを表す非類似度指標となる。従って $\lambda(Z(\mathbf{x}, r))$ の値を \mathbf{x} の関数として図化することで、空間的な混合度の状態を把握することができる。なお円の半径 r は混合度を評価する空間解像度を規定することから、様々な値を試みることで様々な解像度での評価が可能となる。

円の内外における点種 i の構成の差異は、 q_i に対する p_i の比を取ることで把握できる。その比が 1 より大きければ円内での点種 i の割合が相対的に高く、小さければ相対的に低いことになる。比が最大の点種を異なる色相で、 $\lambda(Z(\mathbf{x}, r))$ の値を明度或いは彩度で可視化することで、空間的な混合度をより詳細に把握できる。

次に、対数尤度 $\lambda(Z(\mathbf{x}, r))$ を用いて、内部での点分布構成が外部とは統計的に有意に異なる地域を抽出する。まず、各位置 \mathbf{x} において、 r を 0 から予め定めた値 r_{\max} まで徐々に大きくし、その中で最大の $\lambda(Z(\mathbf{x}, r))$

を与える円を記録、そのときの $\lambda(Z(\mathbf{x}, r))$ の値を $\lambda_{\max}(Z(\mathbf{x}, r))$ とする。この手続きを多数の位置 \mathbf{x} において実施し、得られた $\lambda_{\max}(Z(\mathbf{x}, r))$ を降順に並べ替え、上位の地点の円から順に、互いに重ならないものを規定の個数まで採択する。得られた円の集合について、多重検定の問題を回避するために、FDR (False Discovery Rate) を制御しながら統計的に有意な円を抽出する。

3. 実証分析

本節では、前節で開発した手法の有効性を検証するために、渋谷区における商業施設分布の分析を行う。1995年と2017年の2時点について、商業施設を物販、飲食、サービスの3つに分類し、それぞれの空間分布の混合状態を見る。

図1は1995年における $\lambda(Z(\mathbf{x}, r))$ の分布である ($r=500\text{m}$)。色が濃いほど値の大きいことを表し、緑、赤、青はそれぞれ、物販、飲食、サービスがその地域において特にその割合が高いことを示す。物販は原宿駅東や代官山駅北、飲食は渋谷駅周辺や恵比寿駅南、サービスは新宿駅南でそれぞれその割合が相対的に高くなっていることが分かる。

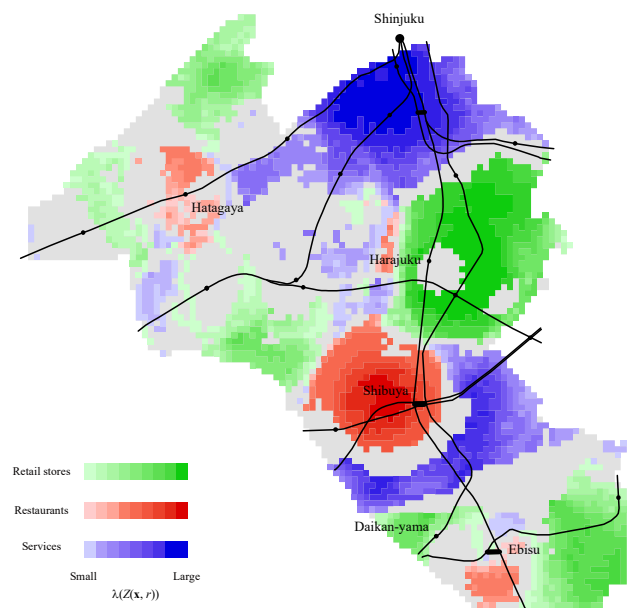


図1 1995年における $\lambda(Z(\mathbf{x}, r))$ の分布 ($r=500\text{m}$)。

図2は2017年における $\lambda(Z(\mathbf{x}, r))$ の分布である。全体的には図1と類似しているが、原宿駅周辺で物販の割合が、恵比寿駅周辺では飲食の割合がそれぞれ高まっていることがわかる。後者は、恵比寿駅周辺における商業集積の拡大によるものと考えられる。

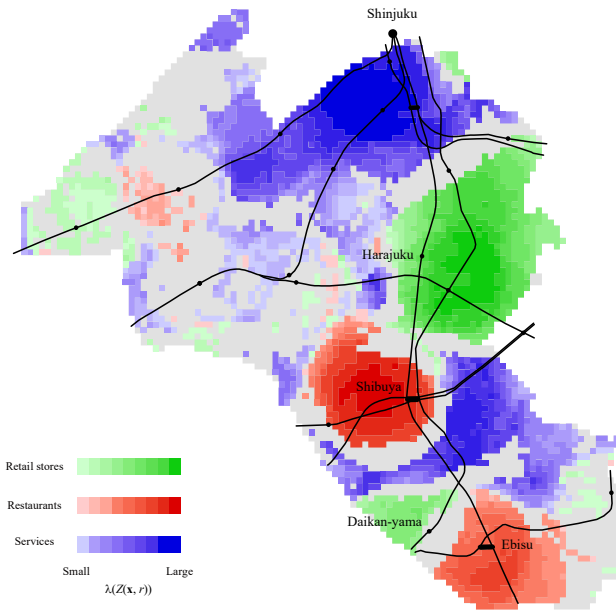


図2 2007年における $\lambda(Z(\mathbf{x}, r))$ の分布 ($r=500m$)。

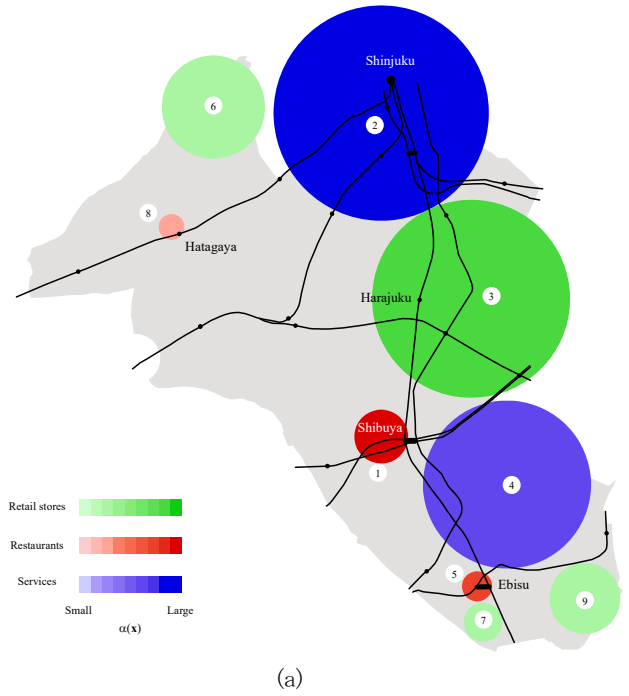
図3は、業種構成が有意に他と異なる地域である。明度が構成の差異の程度、色相が相対的に割合の最も高い業種をそれぞれ表す。図3からは、図1及び2からは読み取ることのできないことがいくつか分かる。例えば図3の渋谷駅周辺の赤円は、図1と比べて遙かに小さく、飲食店の集積が狭い地域に特に集中しているものと考えられる。また、図3の恵比寿駅西側の円の有意性は、図1の印象とは異なって非常に高い。他方、図1における代官山駅北の物販の卓越は、図3aにあるように統計的には有意ではない。

4. おわりに

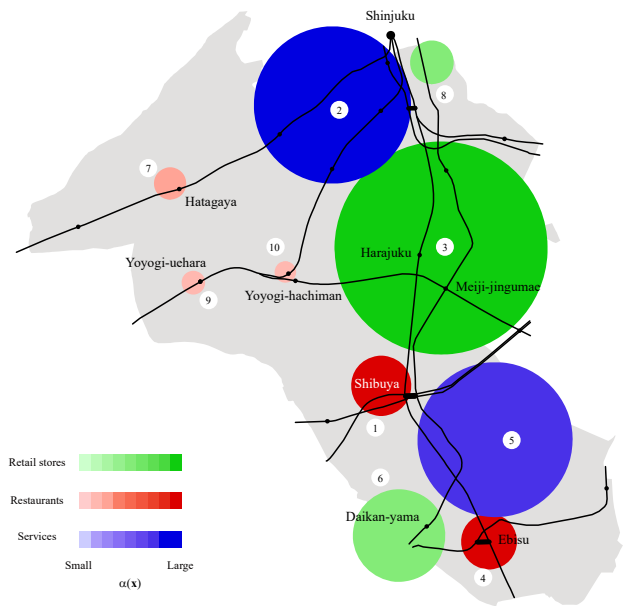
本研究ではNTTタウンページ(株)からデータをお借りしました。ここに記して謝意を表します。

参考文献

[1] Sadahiro, Y. 2019. Statistical analysis of spatial segregation of points. *Computers, Environment*



(a)



(b)

図3 業種構成が有意に他と異なる地域。(a) 1995年, (b) 2017年。