

# 未知パラメータに対する空間的自己相関の検定 - 推定誤差を考慮した Moran's I の分布 -

西颯人\*・山田育穂\*\*

## Spatial Autocorrelation Test for Unknown Parameter - The distribution of Moran's I statistics under the different estimation accuracy -

Hayato Nishi\*, Ikuho Yamada\*\*

This article provides how we can construct the spatial autocorrelation test with Moran's I statistics under the regional difference of observation accuracy. Spatially aggregated observations such as local tax expenditure per capita are noisier in smaller populations. Therefore, the regions where populations are different have different reliability of observation. We treat these observations as estimators of the unknown parameters to adjust the differences in reliability. This treatment enables us to evaluate the reliability quantitatively as the estimation accuracy. The simulation study reveals that the conventional permutation test provides non-uniform type I error rates when populations are not uniform, whereas our adjustment reduced the effect of regional differences in populations. Our adjustment method is applied to the spatial autocorrelation analysis of the local tax expenditure per capita in the Tokyo area.

**Keywords:** 空間的自己相関 (Spatial autocorrelation), Moran's I, 仮説検定 (Hypothesis testing), 偽陽性 (False positive), 推定誤差 (Estimation accuracy)

### 1. はじめに

社会調査等で得られた量的データは、多くの場合自治体単位などに集計され、公開される。例えば内閣府と経済産業省が提供している地域経済分析システム RESAS(<https://resas.go.jp/>)では、2022年8月現在「地方財政マップ」として自治体ごとの一人当たり地方税のヒートマップを表示することができる。こうしたデータの空間分析の際、広く用いられる手法に Moran's I 統計量がある。Moran's I 統計量では空間的自己相関、つまり「全体として、近くの地域で似たような値が観測される傾向にあるか」や「ある地域の近くで似た値が観測される傾向にあるか」を検定することができる。

しかしながら、社会調査の集計単位は一様ではないため、観測値の信頼性に地域差がある可能性が考えられる。例えば、人口一人当たりの税負担額を調査したとする。このとき人口が少ない地域では、個々の税負担額のぶれや観測誤差によって、全体の平

均税負担額が大きな影響を受ける。一方で人口が多い地域であれば、こうした誤差は集計時に統計的には相殺されるため、調査結果として得られた集計値は誤差が小さく信頼性が高いと思われる。

このような信頼性の地域差は、空間的自己相関の検定にも影響を及ぼすと考えられる。人口が少ない地域では極端な値が偶然観測される可能性が、人口の多い地域と比べて高い。このため人口の少ない地域が空間的に集まっていると、偶然に極端な値ばかりが集積する可能性も高くなる。観測値の信頼性の地域差を考慮しなかった場合、こうした偶然の集積を真の集積としてする誤検出が起りやすいと予想され、その結果、検定の信頼性にも地域差が生じることになる。

そこで本研究では、観測値の信頼性の地域差を統制して空間的自己相関を検定する方法を示す。このために、平均税負担額などの集計平均値は真の平均値という未知パラメータの推定量と見なせることを

\* 正会員 一橋大学ソーシャル・データサイエンス教育研究推進センター (Hitotsubashi University)

〒186-8601 東京都国立市中2-1 E-mail : h.nishi@r.hit-u.ac.jp

\*\* 正会員 東京大学空間情報科学研究センター (the University of Tokyo)

利用する。集計平均値が真の平均値の推定量であると考えると、その推定誤差は地域ごとのサンプルサイズから求めることができる。

Moran's I 統計量の分布に関しては、Bivand & Wong (2018)が統計処理のパッケージによって報告される結果が異なることを指摘しており、Sauer et al. (2021)は Moran's I 統計量の分布を推論するための帰無仮説が実装によって異なることがこの原因であるとしている。例えば、Moran's I の分布をつくるために一般的に用いられる並べ替え検定(permutation test)の枠組みを考えよう。この際、並べ替えによって得られる分布の平均と分散は解析的にわかっているが(Anselin, 1995; Sokal et al., 1998)、その分散は完全な並べ替え(total randomization)を仮定した場合と条件付きの並べ替え(conditional randomization)の場合で異なる。このような先行研究から、帰無仮説の設定は Moran's I の分布とそれをを用いた検定結果に影響を与えることがわかる。ここから、前述した信頼性の地域差が帰無仮説に反映されているかが、検定の性質に影響する可能性が考えられる。

Moran's I の分布を推論する際に設定される帰無仮説としては、通常は観測値の交換可能性の仮説に基づく並べ替え検定の枠組みか、正規性の仮定のいずれかが採用される。並べ替え検定は正規性の仮定を要求しない一方で、並べ替え対象となる各地域の等価性が仮定される。しかし人口に地域差がある場合には、観測値の信頼性にも地域差があるので、並べ替え検定の仮説は妥当ではない。一方で正規性の仮定を採用した場合には、推定精度の地域差を考慮した Moran's I の分布を理論的に評価することができる。

本研究では、シミュレーションを通して推定誤差の統制が Moran's I による空間的自己相関の検定結果にどのような影響を与えるのかを明らかにする。また提案手法を用いて、一人当たり地方税額の空間的自己相関を検定する。

## 2. 理論

本章では、推定誤差を考慮した Moran's I 分布の構成方法を示す。

### 2.1. Moran's I 統計量

Tiefelsdorf & Boots (1997)に従い、Moran's I 統計量を定義する。地域 $i(= 1, \dots, n)$ での観測値を $x_i$ と、観測値のベクトルを $\mathbf{x} = (x_1, \dots, x_n)^T$ とする。既知の空間隣接行列を $\mathbf{C}$ として、その $i, j$ 成分を $c_{i,j}$ とする。2つの地域 $i, j$ が隣接している場合 $c_{i,j} = 1$ となり、そうでない場合は0となる。また対角成分について $c_{i,i} = 0$ とする。さらに、 $\mathbf{C}$ を列方向に基準化した重み行列を $\mathbf{W}$ として、その $i, j$ 成分を $w_{i,j}$ とする。以降のシミュレーションでは、隣接関係は Queen 形式の近傍定義を用いて設定する。

このとき、ベクトル $\mathbf{x}$ が観測された対象地域全体における空間的自己相関を示す(Global) Moran's I 統計量 $I(\mathbf{x})$ は以下で定義される。

$$I(\mathbf{x}) = \frac{n\mathbf{x}^T\mathbf{M}\mathbf{W}\mathbf{M}\mathbf{x}}{W_0\mathbf{x}^T\mathbf{M}\mathbf{x}} \quad (1)$$

ただし、 $W_0 = \sum_i \sum_j w_{i,j}$ であり、 $\mathbf{M} = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ と定義した。なお $\mathbf{I}$ は $n \times n$ の単位行列であり、 $\mathbf{1}$ はすべての成分が1の $n$ 次元ベクトル( $n \times 1$ 行列)である。ここで、 $\mathbf{W}$ の代わりに実対称行列 $(\mathbf{W} + \mathbf{W}^T)/2$ を用いても得られる Moran's I の値は変わらない。定義により、ベクトル $\mathbf{M}\mathbf{x}$ の第 $i$ 成分は $x_i$ から標本平均 $\bar{x} = \frac{1}{n}\sum_i x_i$ を引いた値になるので、式(1)は行列を使わずに以下のように表記することもできる。

$$I(\mathbf{x}) = \frac{n\sum_i \sum_j w_{i,j}(x_i - \bar{x})(x_j - \bar{x})}{W_0 \sum_i (x_i - \bar{x})^2} \quad (2)$$

(変形の際は $\mathbf{M}\mathbf{M} = \mathbf{M}$ となることに注意せよ。)

一方で、地域 $i$ 周辺の局所的な空間的自己相関を示す Local Moran's I 統計量 $I_i(\mathbf{x})$ は、隣接行列 $\mathbf{C}$ の $i$ 行目と $i$ 列目を除くすべての要素をゼロに置換した行列 $\mathbf{C}_i$ を $\mathbf{W}$ の代わりに用いて(1)を計算することで求めることができる。これを行列を用いずに表記すると以下ようになる。

$$I_i(\mathbf{x}) = \frac{n(x_i - \bar{x})\sum_j c_{i,j}(x_j - \bar{x})}{2(\sum_j c_{i,j})\sum_j (x_j - \bar{x})^2} \quad (3)$$

なお、上に挙げた式の定数倍が定義となっている場合もあるが、その場合も Moran's I の分布は以降の記述と同様の方法で評価できる。

### 2.2. 推定量としての観測値の誤差

本節では、観測値 $x_i$ を未知のパラメータ $\mu_i$ の推定

量であると考え、その推定誤差を評価できることを示す。

地域*i*で観測された一人当たり地方税負担額を $x_i$ とする。測定上のノイズなどを除外した真の平均地方税負担額を $\mu_i$ として、地域*i*内の個人*k*が負担する地方税額を $y_{i,k}$ とすると、これには個人差や観測誤差があるので、 $y_{i,k}$ は平均 $\mu_i$ 、分散 $\sigma^2$ の正規分布に従うと仮定する。このとき、地域*i*の人口を $m_i$ とすると $x_i = \frac{1}{m_i} \sum_k y_{i,k}$ であるから、 $x_i$ を平均 $\mu_i$ 、分散 $\sigma^2/m_i$ の正規分布に従う $\mu_i$ の推定量とみなすことができる。

よって、平均ベクトルを $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^\top$ 、 $\boldsymbol{\Sigma}$ を(*i, i*)成分が $1/m_i$ である対角行列とすると、観測ベクトル $\mathbf{x}$ は平均 $\boldsymbol{\mu}$ 、分散共分散行列 $\sigma^2 \boldsymbol{\Sigma}$ の多変量正規分布に従うことがわかる。

ここで、平均の空間一様性を仮定した場合、Moran's Iの分布がパラメータに依存しないという興味深い結果を示す。平均の一様性から $\forall i, \mu_i = \mu$ とすると、 $\boldsymbol{\mu} = \mu \mathbf{1}$ と表記できる。このとき、 $\boldsymbol{\Sigma}$ をコレスキー分解などで $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^\top$ と分解できる場合、 $\mathbf{x}$ は各要素が標準正規分布に従うベクトル $\boldsymbol{\varepsilon}$ を用いて

$$\mathbf{x} = \mu \mathbf{1} + \sigma \mathbf{L}\boldsymbol{\varepsilon}$$

と表記できる。これを式(1)に代入すると、

$$\begin{aligned} I(\mathbf{x}) &= \frac{n\mathbf{x}^\top \mathbf{M}\mathbf{W}\mathbf{M}\mathbf{x}}{W_0 \mathbf{x}^\top \mathbf{M}\mathbf{x}} \\ &= \frac{n\boldsymbol{\varepsilon}^\top \mathbf{L}^\top \mathbf{M}\mathbf{W}\mathbf{M}\mathbf{L}\boldsymbol{\varepsilon}}{W_0 \boldsymbol{\varepsilon}^\top \mathbf{L}^\top \mathbf{M}\mathbf{L}\boldsymbol{\varepsilon}} \end{aligned} \quad (4)$$

が得られる。ここで、 $\mathbf{M}\mathbf{L} = \mathbf{0}$ 、 $\mathbf{M}\mathbf{x} = \mu \mathbf{M}\mathbf{1} + \sigma \mathbf{M}\mathbf{L}\boldsymbol{\varepsilon} = \sigma \mathbf{M}\mathbf{L}\boldsymbol{\varepsilon}$ を用いた ( $\mathbf{0}$ はゼロベクトル)。式(4)では、平均パラメータ $\mu$ は行列 $\mathbf{M}$ を乗算することによって、分散パラメータ $\sigma^2$ は分子と分母の相殺によって、最終的な式には含まれていない。すなわち、共分散構造を示す $\boldsymbol{\Sigma}$ が分かっているならば、Moran's I 統計量は未知パラメータ $(\mu, \sigma^2)$ に依存しない不動統計量(pivotal statistics)となる。

### 2.3. 正規性を仮定した場合の Moran's I の分布

$\mathbf{x}$ が正規分布に従う場合の $I(\mathbf{x})$ の分布は、Tiefelsdorf (1998)等で、近似に関しては Tiefelsdorf (2002)で示されている。実際の計算に関しては Bivand et al. (2009)が詳しい。

$I(\mathbf{x})$ がある値 $I_{obs}$ 以下となる確率 $\Pr[I(\mathbf{x}) \leq I_{obs}]$ は、式(4)の変形によって以下と等価であることがわかる。

$$\Pr[I(\mathbf{x}) \leq I_{obs}]$$

$$= \Pr[\boldsymbol{\varepsilon}^\top (n\mathbf{L}^\top \mathbf{M}\mathbf{W}\mathbf{M}\mathbf{L} - I_{obs} W_0 \mathbf{L}^\top \mathbf{M}\mathbf{L}) \boldsymbol{\varepsilon} \leq 0] \quad (5)$$

ここで、 $\mathbf{T} = n\mathbf{L}^\top \mathbf{M}\mathbf{W}\mathbf{M}\mathbf{L} - I_{obs} W_0 \mathbf{L}^\top \mathbf{M}\mathbf{L}$ とにおいて、その固有値分解を $\mathbf{T} = \mathbf{E}^\top \boldsymbol{\Lambda} \mathbf{E}$ とする(ここで $\mathbf{W}$ の代わりに実対称行列 $(\mathbf{W} + \mathbf{W}^\top)/2$ を用いると、固有値分解可能であることが保証される)。 $\boldsymbol{\Lambda}$ は固有値 $(\lambda_1, \dots, \lambda_n)$ を対角成分に並べた対角行列である。 $\mathbf{E}$ が直交行列となるようにすれば、 $\mathbf{z} = \mathbf{E}\boldsymbol{\varepsilon}$ も独立な標準正規分布に従うことから、式(5)内の不等号の左辺

$$\boldsymbol{\varepsilon}^\top \mathbf{C} \boldsymbol{\varepsilon} = \sum_i \lambda_i z_i^2 \quad (6)$$

は一般化カイ二乗分布に従う。この変形に関しては Das & Geisler (2021)が詳しい。以上から、一般化カイ二乗分布の累積分布を調べることで、未知パラメータ $(\mu, \sigma^2)$ を用いずに Moran's I 統計量の累積分布を評価することができる。なお、Local Moran's I に関しても、Global Moran's I と同様に式(1)の形式で表現できることから、同様の方法で累積分布を評価することができる。

### 2.4. 推定精度を補正した検定

以上の性質から、次のようなステップによって $x_i$ の信頼性を考慮した空間的自己相関の検定を行うことができる。

1. 式(1)を用いて Moran's I 統計量を計算し、これを  $I_{obs}$  とする
2. 各地域の人口 $(m_1, \dots, m_n)$ を用いて共分散構造 $\boldsymbol{\Sigma}$ を作成することで、推定精度を評価する(第2.2節参照)
3. 式(5)を用いて $\Pr[I(\mathbf{x}) \leq I_{obs}]$ を評価し、検定する  
最後のステップでは、両側検定と片側検定のどちらでも標準的な計算方法に従って実施することができる。この検定は第2.2節で示したデータ生成過程のモデルに基づいていることから、この検定を Model-based Test と呼ぶことにする。なお、並べ替えによる Moran's I の検定は Randomization Test や Monte-Carlo Test と呼ばれる場合も多いが、本稿では Moran's I の分布を評価するための帰無仮説の違いが本質的であるため、仮説を直接表現した並べ替え検

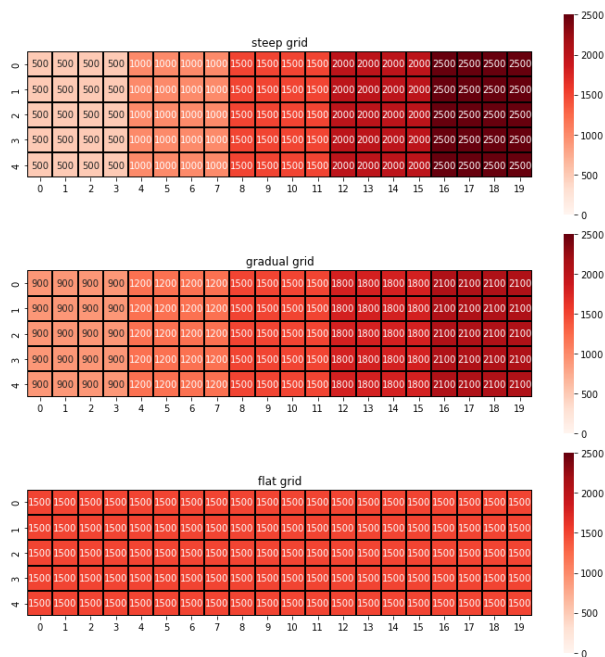


図1：3パターンの人口分布. セル内の数字はそのセルの人口 $m_i$ を表す.

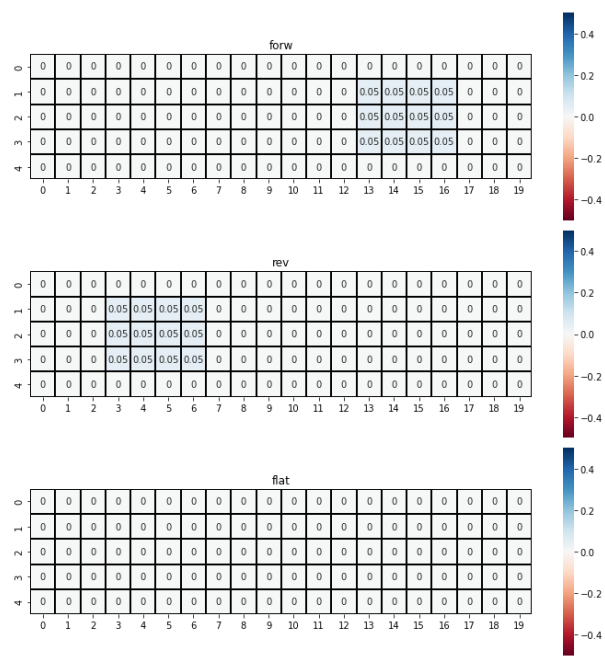


図2：3パターンの平均値分布. セル内の数字はそのセルの平均値 $\mu_i$ を示す.

定(Permutation Test)という呼称で統一することにする.

### 3. シミュレーションによる評価

本章ではシミュレーションを行うことで, Model-based Test のふるまいを並べ替え検定と比較する. シミュレーションの設定は山田(2020a, 2020b)を参考にした.

#### 3.1. 方法

20×5 個のセルに分割された長方形の地域を想定し, そこから観測値をランダムに生成するシミュレーションを行う. 具体的には以下のようなデータ生成過程(DGP: Data Generation Process)を設定した.

各セルには人口 $m_i$ , 平均 $\mu_i$ , 分散 $\sigma^2$ が割り当てられており, 住民はそれぞれ平均 $\mu_i$ , 分散 $\sigma^2$ の正規分布に従う値(税負担額などを想定している)を持つとする. このとき, 各住民の持つ値の平均値がセル $i$ の観測値 $x_i$ として得られるとすると, これをシミュレーションするには平均 $\mu_i$ , 分散 $\sigma^2/m_i$ の正規分布に従う擬似乱数を生成して $x_i$ とすればよい. ただし, 分散はすべての地域で共通して $\sigma^2 = 1.0$ とした.

地域間で人口の偏りがある場合を想定して, 図1

で示す3通りの人口パターンを設定した. flat grid (図1下)は人口分布が一様なパターンで, gradual grid (図1中央)は右に行くほど徐々に人口が多くなるパターンである. steep grid (図1上)は gradual grid よりも人口の偏りが大きいパターンである. なお, 3つの人口パターンは人口の総計が同じになるように設定されている.

また, クラスターの有無や所在地に関して, 図2で示す3パターンの平均値 $\mu_i$ の分布を設定した. forw (図2上)は右側にクラスターがある設定であり, 人口の多い場所にクラスターが存在する. 一方で rev (図2中央)は左側にクラスターが存在し, 人口の少ないところにクラスターがある設定である. flat (図2下)はクラスターがなく, すべての地域が同じ平均値を持っている設定となっている. よって, クラスターパターンが flat のときに有意な検定結果が出た割合が, 偽陽性率となる.

以上の3×3=9通りの設定のもとでシミュレーションを繰り返し行い, Model-based Test と Permutation Test による両側検定の結果を比較する. 観測値の生成は1,000回行い, 5%有意水準で有意な結果が出た割合を記録した. また Permutation Test に

表 1 : Global Permutation Test で有意な結果が出た

クラスター パターン	割合		
	人口パターン		
	steep grid	gradual grid	flat grid
forw	0.559	0.675	0.703
rev	0.517	0.642	0.716
flat	0.104	0.065	0.046

においては 999 回の並べ替えを行った。なお Local Moran's I の検定においては、すべての観測値を並べ替える Total Permutation と、対象となる地域*i*以外での観測値を並べ替える Conditional Permutation とで 2通りの方法があるが、本研究ではより一般的と思われる Conditional Permutation によって並べ替え分布を生成した。

### 3.2. Global Moran's I の結果

表 1, 2 はそれぞれ Permutation Test, Model-based Test によって 5%有意な結果が出た割合を示す。クラスターパターン forw, rev の場合はクラスターが存在し、flat の場合は存在しないので、表中の数値は前者の場合は検出率、後者の場合は偽陽性率を示す。5%有意水準で検定を行っているため、偽陽性率は 5%以下となっていることが望ましい。

Permutation Test の場合、人口に偏りのない flat grid のときは偽陽性率が 5%程度になっているものの、人口に偏りが大きいほど偽陽性率が大きくなっていることがわかる。一方で Model-based Test の場合は、どの人口パターンでも偽陽性率は 5%程度となっている。この結果から 5%有意水準の仮説検定としては、偽陽性率を適切にコントロールできている Model-based Test の方が望ましいと言える。

一方で検出率を見ると、Permutation Test の方が全体的に高めの値となっている。この結果は、偽陽性率を適切に抑制した結果、Model-based Test はやや保守的な検定となったと解釈できる。またいずれの検定方法を用いても、人口の偏りが少ない方が検出率は高くなる傾向が見られた。

### 3.3. Local Moran's I の結果

各セルが Local Moran's I 統計量により 5%有意とされた割合を図 3, 4 に示す。ここでは簡単のため

表 2 : Global Model-based Test で有意な結果が出た

クラスター パターン	割合		
	人口パターン		
	steep grid	gradual grid	flat grid
forw	0.421	0.632	0.697
rev	0.408	0.618	0.716
flat	0.049	0.057	0.047

に、多重比較補正は行っていない。

まず帰無仮説が正しくクラスターが存在しない flat (図 3, 4 の下段にあるパネル) の結果を見て、偽陽性率が 5%以下にコントロールできているかを確認する。Permutation Test (図 3) の場合は、人口分布が一樣な場合は偽陽性率が全地域でおおむね 5%である一方で、人口に偏りがある場合は人口が少ない左側のセルで偽陽性率が最大 15%程度までに高くなっている。一方で Model-based Test (図 4) では、人口分布の偏りに関わらずにどのセルでも 5%程度の偽陽性率となっており、すべてのセルで公平な検定となっていることがわかる。

一方でクラスターが存在する場合には、どちらの手法を用いても平均値が大きな値となっているセルおよびその近傍で検出が多くなっていることが確認できる。クラスターの検出率自体は Permutation Test の方がわずかに高いが、クラスターのないセルでの「検出」を見ると、こちらの場合でも Permutation Test は左側の人口の少ないセルで誤検出が多くなっている。Model-based Test の場合はクラスターのないセルでは人口に関わらず偽陽性率がおおむね 5%以下となっているため、偽陽性率のコントロールが適切に行われていると言える。

以上の結果から Model-based Test は、人口の偏りがあっても設定した有意水準に即した検定を行うことができるという意味で、望ましい性質を持つ検定である。

## 4. 一人当たり地方税の分析

Model-based Test を用いて、東京都（諸島部を除く）・埼玉県・神奈川県 の市区町村別一人当たり地方税の分析を行う。Model-based Test を用いることで、

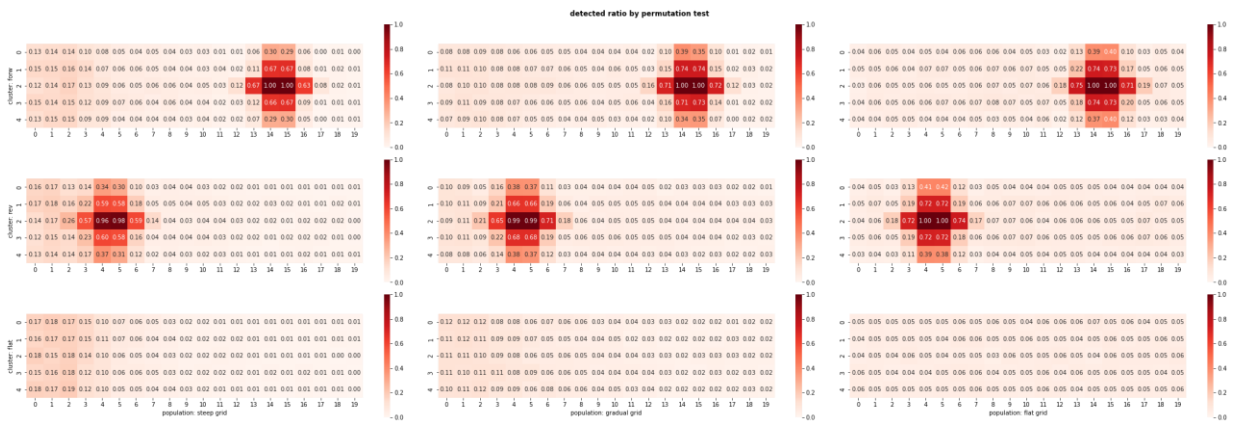


図3：Local Permutation Test で5%有意と検出された割合

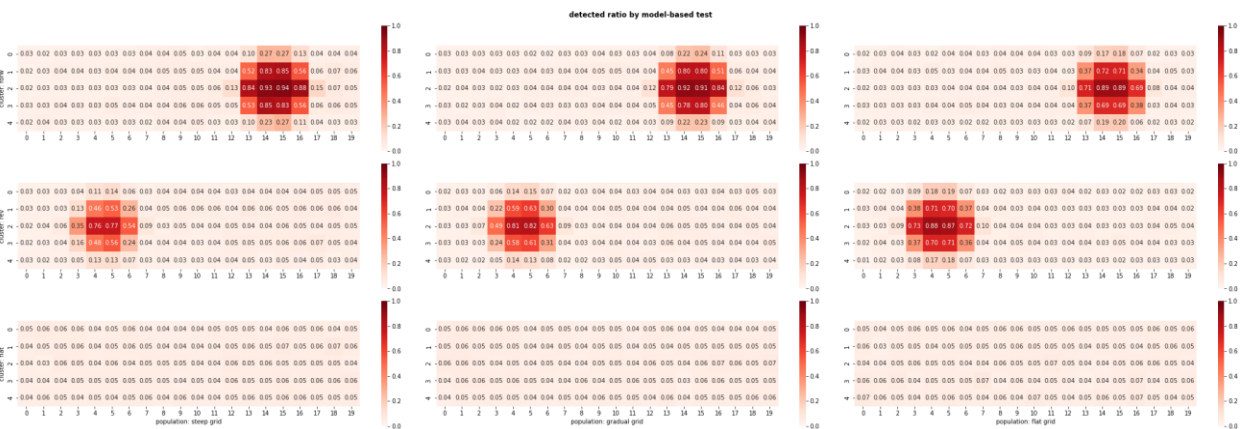


図4：Local Model-based Test で5%有意と検出された割合

自治体ごとの人口の差を考慮した検定を行うことができる。

#### 4.1. データ

人口及び地方税のデータは RESAS と同様に総務省地方財政状況調査関係資料の「令和2年度市町村別決算状況調」を使用して一人当たり地方税額を算出した。また行政界ポリゴンデータは国土数値情報から取得した。隣接行列は PySAL (<https://pysal.org/>) version 2.1.0 を用いて Queen 形式で隣接関係を判定して作成した。

図5に人口分布を示した。また図6は一人当たり地方税のヒートマップである。なお図6中の赤破線は、隣接関係にある自治体を結んでいる。

#### 4.2. 結果

一人当たり地方税の Global Moran's I 統計量は 0.163 であり、Model-based Test (両側検定) による p 値は 0.009 であった。このことから、一人当たり地

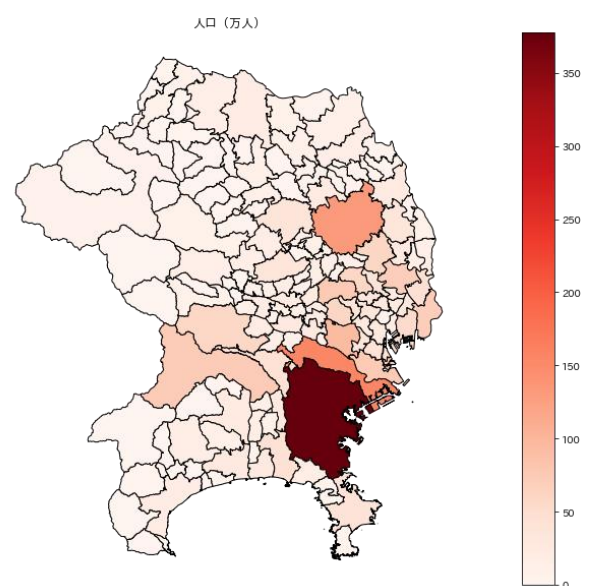


図5：人口（万人単位）の空間分布

方税には有意水準 5%で統計的に有意な正の空間的自己相関があり、近傍にある自治体では一人当たり



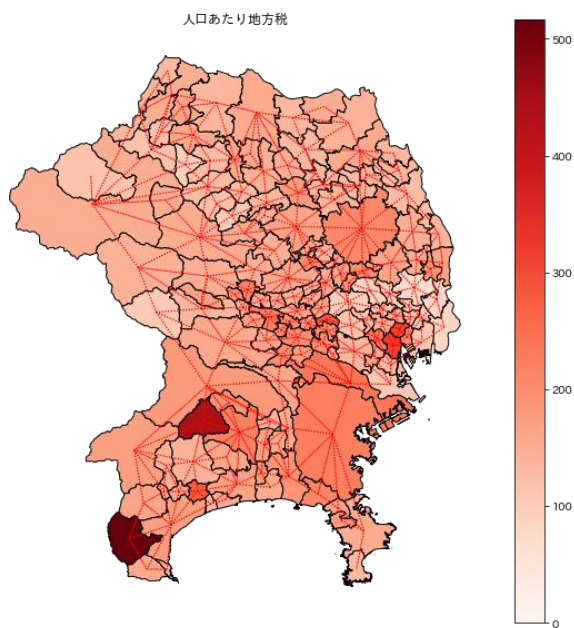


図6：人口あたり地方税の空間分布

地方税が類似する傾向にあると解釈できる。

図7には Local Moran's I の分布を示した。両側検定により 5%有意と判断された自治体の枠線は黒の太線で示している。ほとんどの自治体で正の空間的自己相関が確認されたが、有意であったのは主に東京23区と横浜市近辺の東京湾沿いの自治体であった。また内陸でも人口の多い自治体やその近辺では有意となっている場合もあった。図6と比較すると、東京23区は地方税が安いクラスタとなっており、一方で横浜市付近や神奈川県西部は地方税が高いクラスタになっていると考えられる。図6では箱根町（南西端）と清川村（県西部）の地方税額が目立って高いが、清川村近辺ではクラスタが有意となっている一方で、箱根町では有意ではない。これは、人口の少ない自治体では集計値が不安定性になるため、これを考慮すると統計的に有意な空間的自己相関ではないと判断したためと考えられる。

## 5. おわりに

本研究では、Moran's I 統計量を用いた空間的自己相関の検定を行う際に、人口の多寡による集計平均値の信頼性をコントロールする方法を示した。Moran's I による検定は観測値の並べ替えに基づいて行われることが多いが、人口の多寡がある場合は集

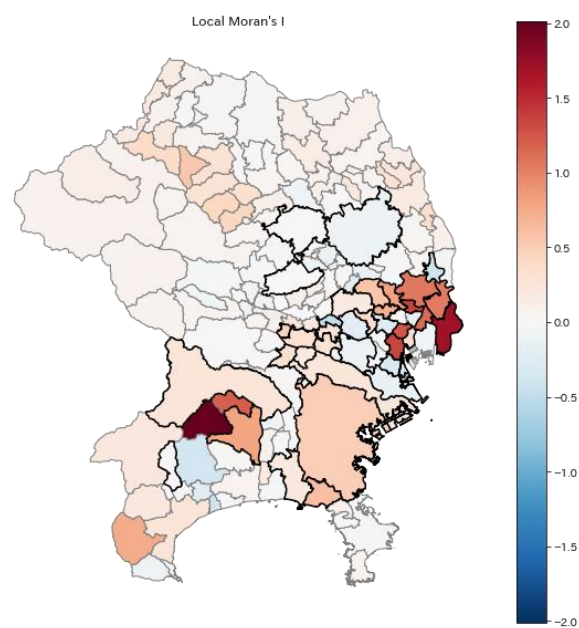


図7：Local Moran's I 統計量の空間分布。5%有意な自治体の境界は黒太線にした。

計平均値の信頼性に域差が生じる。そこで集計平均値を未知パラメータの推定量とみなし、正規性の仮定を導入することによって、信頼性の地域差を考慮した Moran's I の分布を導出できることを示した。

シミュレーションにより、以下のようなことが示唆された。1)人口に地域差がある場合、並べ替え検定では人口が少ない地域での偽陽性率が設定した有意水準以上に増大する場合がある。2)提案手法は並べ替え検定に比べてやや保守的である代わりに、人口に地域差があっても偽陽性率はおおむね有意水準以下に維持される。3)いずれの検定方法でも、人口の偏りが小さい方がクラスタの検出はしやすい傾向にある。

さらに提案手法を東京都・埼玉県・神奈川県の一人口あたり地方税の分析し、人口の地域差を考慮した分析結果を示した。まず Global Moran's I を用いると、統計的に有意な正の空間的自己相関が見いだされた。一方で Local Moran's I を用いると、東京湾に近い自治体や人口の多い自治体や統計的に有意な関係が見いだされたが、その他の地域の空間的自己相関は統計的に有意ではなかった。この原因として、人口が少ない自治体では集計値が相対的に大きければつきを持つということが考えられる。

本研究では正規分布に従う観測値を想定した検定方法を示した。この方法は DGP のパラメータに関する事前知識を要求しないため応用可能性が高いと考えられるが、実際の社会調査等では件数データなどの正規分布に従わない観測値も多く存在する。このような件数データや割合データへの拡張は、今後の研究課題としたい。

## 謝辞

本研究は JSPS 科研費 22H00245 の助成を受けたものです。

## 参考文献

- 山田育穂 (2020a). 標本に基づく空間データの利用が空間解析に及ぼす影響についての考察 空間的自己相関の解析を題材に. 「都市計画論文集」, **55**(3), 435-442.
- 山田育穂 (2020b). 空間解析における標本に基づく空間データの利用についての一考察. 『あいまいな時空間情報の分析』(浅見泰司・薄井宏行編著), 古今書院, 19-37.
- Anselin, L. (1995). Local Indicators of Spatial Association-LISA. *Geographical Analysis*, **27**(2), 93-115.
- Bivand, R., Müller, W. G., & Reder, M. (2009). Power calculations for global and local Moran's I. *Computational Statistics & Data Analysis*, **53**(8), 2859-2872.
- Bivand, R. S., & Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial association. *TEST*, **27**(3), 716-748.
- Das, A., & Geisler, W. S. (2021). A method to integrate and classify normal distributions. *Journal of Vision*, **21**(10), 1.
- Sauer, J., Oshan, T., Rey, S., & Wolf, L. J. (2021). The Importance of Null Hypotheses: Understanding Differences in Local Moran's under Heteroskedasticity. *Geographical Analysis*, Early View.
- Sokal, R. R., Oden, N. L., & Thomson, B. A. (1998). Local Spatial Autocorrelation in a Biological Model. *Geographical Analysis*, **30**(4), 331-354.
- Tiefelsdorf, M., & Boots, B. (1997). A Note on the Extremities of Local Moran's  $I_i$ s and Their Impact on Global Moran's I. *Geographical Analysis*, **29**(3), 248-257.
- Tiefelsdorf, M. (1998). Some practical applications of Moran's  $I$ 's exact conditional distribution. *Papers in Regional Science*, **77**(2), 101-129.
- Tiefelsdorf, M. (2002). The saddlepoint approximation of Moran's  $I$ 's and local Moran's  $I_i$ 's reference distributions and their numerical evaluation. *Geographical Analysis*, **34**(3), 187-206.