

生活利便施設へのアクセシビリティが人口分布の規定要因と考える統計的分析方法

A statistical method for determining whether the population distribution in a region is significantly affected by the accessibility to daily facilities and whether population is large at places where facilities are close by

増山 篤*

Atsushi Masuyama

This paper proposes a statistical method for determining whether the population distribution in a region is significantly affected by the accessibility to daily facilities and whether population is large at places where facilities are close by. In urban planning, it has been of interest to explore the relationships between population distribution and other distributions. In particular, in recent years, increasing attention is paid to whether a population distribution is affected by the accessibility to daily facilities such as medical and commercial facilities, and also to whether people tend to live close to such facilities and therefore population density is high around the facilities. Taking into account the fact that there is generally positive spatial autocorrelation in population density, this paper proposes a statistical method for investigating these questions, and, using the method, empirically analyzes the distributions of population and supermarkets in Hirosaki-shi, Aomori.

Keywords: population distribution, accessibility, daily facilities, spatial autocorrelation, statistical method
人口分布、アクセシビリティ、生活利便施設、空間的自己相関、統計的分析方法

1. はじめに

都市計画分野では、その時々時代の背景を受け、人口またはその増減の空間分布を分析する研究が行われてきた。「人口またはその増減の空間分布」と書いたが、これを「人口指標分布」と言うことにしよう(その定義より、人口指標分布は、いわゆる「人口分布」を含む包括的な概念である)。人口指標分布を分析する研究の多くは、単に、どこにどの程度の人口があり、どこでどの程度増減しているかを記述するにとどまらない。人口指標分布が、何かを要因として規定されているか、あるいは、何かに影響を及ぼしているか、さらには、何かと相互に関連し合っているかに関心を寄せ、その関心に沿った分析を行っている。まとめて言うならば、人口指標分布と他の何かとの関係を探っている。それらの研究を分類すると、おおむね時系列的に並べうる五つのタイプが挙げられる。

第一に、高度経済成長期に都市へ人口が集中し、その結果として様々な都市問題が発生したことを背景とするものが挙げられる^{1), 2), 3), 4), 5), 6), 7)}。これらのうち、大村ら⁴⁾、浅見⁶⁾では、人口指標分布と道路との関係を分析している。浅見⁷⁾では、人口指標分布と“線引き”(の結果たる市街化区域界)との関係を分析している。これらはいずれも、人口指標分布が他の何かを要因として規定されている、あるいは、他の何かと浅からぬ関連を持つ、ということは想定するが、他の何かに影響を及ぼしているということは想定していないように見られる。

第二に、都市中心部における人口減少に関する研究が挙げられる^{8), 9), 10), 11)}。これらは、アンケート等を通じて、研究対象地区における、その地区固有の人口減少要因を探るものである。

第三に、(人口指標分布の一つである)人口分布を所与とみなし、その施設立地への影響を分析するものがある^{12), 13)}。これらは、高度経済成長期も終わり、日本の総人口も定常的狀態に達した後に行われたものである。その時期ゆえ、ごく自然に人口分布を所与とみなす前提が採られているかと思われる。

第四に、近年のコンパクトシティを巡る議論に端を発する研究が挙げられる^{14), 15), 16), 17)}。それらのうち、海道¹⁵⁾では、人口密度と生活利便施設立地状況の相関関係を分析している。この相関分析では、人口分布を所与とし、これが施設立地に影響を与えるという前提に立っていることが伺える。しかし、この研究以外では、人口指標分布を所与と考えず、むしろ、物的施設へのアクセシビリティ(より平易な表現を用いれば、“近づきやすさ”とでもなる)によって影響されるという前提に立っている。森永ら¹⁴⁾では、GISによる視覚化結果から、生活利便施設周辺では人口密度が高くなるように見受けられると述べている。橋本ら¹⁶⁾では、最寄り駅までの距離が近いほど、人口密度の増加がみられるかを分析している。内原・吉川¹⁷⁾では、生活利便性の高い場所において、人口増加がみられるかどうかを分析している。

第五に、近年急速に進む高齢化を背景とする研究が挙げられる^{18), 19)}。これらでは、高齢者人口密度が、他の年齢層と比較して、生活利便施設周辺で顕著に高くなっているかどうか、分析している。現在の高齢者が、加齢とともに、生活利便施設へのアクセシビリティが高い場所を指向する居住選択を行う可能性を念頭に置きつつ、分析を行っている。

このように、最近の研究では、生活利便施設へのアクセシビリティが人口指標分布を規定する要因であり、施

* 正会員・弘前大学人文学部(Hirosaki University, Faculty of Humanities)

設の近くで顕著に人口が多いか、ということが大きな関心となっている。この関心に対して、何か明快な答えを与えられるとすれば、その有力な一つは、仮説検定型のデータ分析方法であろう。以下では、「人口指標分布」のうち、いわゆる「人口分布」(ある一時点における人口の空間的広がり)に対象を絞り、まさにそうした分析方法について考えていこう。

一般的な仮説検定型データ分析の枠組みによって、生活利便施設へのアクセシビリティを要因として人口分布が規定され、施設へ近づきやすい場所で人口が顕著に多いかどうかを判別するとなれば、例えば、以下の四つのステップを経る方法が考えられる。

- Step 1: 人口分布が生活利便施設へのアクセシビリティによって規定されるというモデルを考える。
- Step 2: モデルパラメータを推定し、また、何らかの指標によって、モデルのデータへの適合度を測る。
- Step 3: “他にありえたかもしれない、今のものとは違う人口分布”を生成する確率的プロセスを考え、モデルパラメータやモデル適合度の確率分布を導出する。
- Step 4: この確率分布の下で、推定されたパラメータやモデル適合度の統計的有意性を吟味する。

Step 1の代わりに、生活利便施設へのアクセシビリティによって人口分布が規定されている程度、あるいは、施設の近く場所で人口が多くなっている程度を表す指標を考え、それを検定統計量とする方法も考えられる。この場合、Step 2に代わって、分析対象データから検定統計量を算出することになる。そして、Step 3に代わって、“他にありえたかもしれない、今のものとは違う人口分布”を考えたときの検定統計量の確率分布を導くことになる。

以後の記述の簡便を図り、用語を一つ定義しておく。Step 3における“他にありえたかもしれない、今のものとは違う人口分布”を、以下では、「可能性人口分布」と言うことにする。

先行研究において、上記の四つのステップのうち最初の二つまでを行う方法は提案されている^{20), 21)}。しかし、残る二つまでは実行せず、モデルのデータへの適合度やモデルパラメータの統計的有意性を吟味するものではない。つまり、施設へのアクセシビリティを要因として人口分布が規定され、そして、施設へ近づきやすい場所で顕著に人口が多いか、という問いに答えるものではない。

上記のStep 3および4を実行するためには、可能性人口分布を生み出す確率的プロセスを解析的もしくは数値的に取り扱えるようにする必要がある。さらに、そのためには、可能性人口分布が満たすべき要件を明確にしておく必要がある。そこで、この要件について考え

てみよう。

都市内において、人は集まって住む。このことはごく普遍的に観察される事実である。すると、可能性人口分布も、この事実と反するようなものであってはならない。つまり、このことが、可能性人口分布の満たすべき要件であろう。

この要件を踏まえつつ、可能性人口分布を生み出す確率的プロセスを解析的に扱うことは、いかにも困難である。しかし、後に示すように、一定の空間的自己相関水準を持つ空間分布をシミュレートする方法²²⁾を援用することで、可能性人口分布を数値的に扱うことは可能である。これにより、上記のStep 3および4が実行される。そして、上記の四つのステップが、通して実行される。

この論文では、まず、これまでに提案された分析方法等^{20), 21), 22)}に対し、必要なアレンジを加え、人口分布が生活利便施設へのアクセシビリティによって規定されているか、そして、施設の近くで人口が顕著に多いか、を統計的に判別する方法を提案する。そして、この方法を、青森県弘前市における総人口および高齢者分布、商業・医療施設ポイントデータに対して適用した結果を示す。最後に、この論文の内容をまとめ、今後の課題を述べる。

2. 各種施設へのアクセシビリティによって人口分布が規定されているか統計的に判別する方法

この章では、各種施設へのアクセシビリティによって人口分布が規定されているかどうか、そして、施設の近くで人口が顕著に多いかどうか、を統計的に判別する方法を提案する。

今、 n 個の空間単位 S_1, S_2, \dots, S_n からなる領域 S があるものとする。空間単位 i の面積を A_i と表すことにする。空間単位 S_i の中に適当な代表点を取ったとき、その位置を表すベクトルを \mathbf{x}_i とする。また、空間単位 S_i における人口を P_i とする。さらに、空間単位 S_i における人口密度を p_i とすると、

$$p_i = \frac{P_i}{A_i}, \quad (1)$$

となる。

S 内には、スーパーマーケットであるとか内科系医療機関であるといった、ある特定のタイプの施設が m 個分布しているものとする。施設 j の位置を表すベクトルを \mathbf{y}_j とする。二地点 \mathbf{x} , \mathbf{y} 間の距離を、 $\|\mathbf{x} - \mathbf{y}\|$ と表すことにする。

ある地点 \mathbf{x} から一つの施設 j へのアクセシビリティを考えよう。ごく自然に考えると、地点 \mathbf{x} から施設 j に対するアクセシビリティは、 $\|\mathbf{x} - \mathbf{y}_j\|$ の単調減少関数であろう。このときの関数型にはさまざまなものが考えられるが、最も一般的なものは指数関数であろう。そこで、

この研究でも指数関数を用いることにする。つまり、地点 \mathbf{x} から施設 j に対するアクセシビリティを $C_j(\mathbf{x})$ と表すと、ある正の定数 A が存在し、

$$C_j(\mathbf{x}) = A \exp(-\beta \|\mathbf{x} - \mathbf{y}_j\|), \quad (2)$$

とする。なお、ここで β は正の定数であり、距離に応じてアクセシビリティがどの程度減少するかを決めるパラメータである。以後、このパラメータを、「距離減衰パラメータ」と言うことにする。

次に、施設 j が属するタイプの施設全体に対するアクセシビリティを考えよう。地点 \mathbf{x} から、あるタイプの施設全体へのアクセシビリティを $C(\mathbf{x})$ と書くことにする。ごく普通に発想すると、居住場所の周辺における施設数が多いほど $C(\mathbf{x})$ も大きいと考えられる。このことから、

$$C(\mathbf{x}) = \sum_{j=1}^m C_j(\mathbf{x}), \quad (3)$$

とするのが、一つのごく妥当な考え方であろう。

ここで、施設に対するアクセシビリティによって人口密度が規定されるというモデルを考える。より具体的には、 S 内の任意の地点における人口密度は、その地点におけるアクセシビリティに正比例するというモデルを考える。地点 \mathbf{x} における人口密度を $p(\mathbf{x})$ とすると、このモデルは、

$$p(\mathbf{x}) = kC(\mathbf{x}), \quad (4)$$

と書ける。

$p(\mathbf{x})$ および $C(\mathbf{x})$ を基準化し、以下の式で定義される $p'(\mathbf{x})$ および $C'(\mathbf{x})$ を考えよう。

$$p'(\mathbf{x}) = \frac{p(\mathbf{x})}{\int_S p(\mathbf{x}) d\mathbf{x}}, \quad (5)$$

$$C'(\mathbf{x}) = \frac{C(\mathbf{x})}{\int_S C(\mathbf{x}) d\mathbf{x}}, \quad (6)$$

すると、今考えているモデルは、式(4)よりも簡潔に、

$$p'(\mathbf{x}) = C'(\mathbf{x}), \quad (7)$$

と書ける。なお、 $p(\mathbf{x})$ をもとに構成された $p'(\mathbf{x})$ においては、定数 A が消えており、式(4)で表されるモデルに代わって、式(7)で表されるモデルを考えることで、定数 k が消えていることに注意されたい。

式(4)あるいは式(7)で表されるモデルでは、人口密度は施設へのアクセシビリティのみによって規定されることになるが、無論、現実にはそのようなことはない。アクセシビリティ以外のランダムな要因 $e(\mathbf{x})$ (人々が居住場所を決める際に、施設へのアクセシビリティ以外に考慮

するかもしれない諸々の要因) も加わり²⁰⁾、人口密度は決まっていると考えの方が自然だろう。

アクセシビリティにランダムな要因も加わって人口密度が規定されるというモデルは、

$$p'(\mathbf{x}) = C'(\mathbf{x}) + e(\mathbf{x}), \quad (8)$$

と書ける。なお、ここで

$$E(e(\mathbf{x})) = 0, \quad (9)$$

$$\int_S e(\mathbf{x}) d\mathbf{x} = 0, \quad (10)$$

である ($E(\cdot)$ は、期待値を表すオペレータ)。

式(8)で表されるモデルを「居住モデル」と呼ぶことにしよう。居住モデルを、人口データ、施設位置データにあてはめるには、何らかの意味で $p'(\mathbf{x})$ と $C'(\mathbf{x})$ が近くなる ($e(\mathbf{x})$ が小さくなる) 距離減衰パラメータ β を推定すればよい。そのために、まず $p'(\mathbf{x})$ と $C'(\mathbf{x})$ の近さを測る適当な指標を考えよう。

$p'(\mathbf{x})$ も、 $C'(\mathbf{x})$ も、 S 全体に渡って積分すると 1 となるため、各種の確率密度関数間の距離が、両関数の近さを測る指標として用いられよう。このようなとき、先行研究^{18), 19), 20), 21)} では、カルバック・ライブラー情報量²³⁾ (以下、K-L 情報量)

$$\int_S p'(\mathbf{x}) \log \frac{C'(\mathbf{x})}{p'(\mathbf{x})} d\mathbf{x}, \quad (11)$$

を用いている。K-L 情報量は、0 以上の値を取り、あらゆる \mathbf{x} に対して式(7)が成立するときにおいてのみ、理論的最小値 0 を取るという性質を持つ。この性質はいかにも好ましい。ただし、いずれかの \mathbf{x} において $p'(\mathbf{x}) = 0$ となるとき、K-L 情報量は (対数の真数部分が計算できず) いかに不具合である。そこで、この論文では、K-L 情報量に代わって、以下の式で定義される L_2 ノルム²⁴⁾ を用いることにしよう。

$$\int_S \{p'(\mathbf{x}) - C'(\mathbf{x})\}^2 d\mathbf{x}. \quad (12)$$

L_2 ノルムは、K-L 情報量と同様、0 以上の値を取り、あらゆる \mathbf{x} において式(7)が成立するときにおいてのみ、理論的最小値 0 を取る。

L_2 ノルムによって $e(\mathbf{x})$ の大きさを表すならば、式(12)を最小にする β を求めることで、居住モデルを人口データ、施設位置データに最も適合させる β が推定される。各空間単位 S_i に人口 P_i が与えられ、その位置がベクトル \mathbf{x}_i で表される点によって代表されているとき、その推定を行うための計算手順は以下ようになる。

まず、すべての空間単位に関して足しあわせたとき、その合計が 1 となるような“基準化された”人口密度を

考える。これまで用いてきた記号を使えば、 $p'(\mathbf{x})$ に対応するものである。空間単位 S_i に対応する基準化された人口密度は、

$$\frac{P_i}{A_i \sum_{j=1}^n P_j}, \quad (13)$$

となる。同様に、空間単位 S_i に対応する基準化された（施設への）アクセシビリティは、

$$\frac{\sum_{j=1}^m \exp(-\beta \|\mathbf{x}_i - \mathbf{y}_j\|)}{\sum_{k=1}^n \left\{ A_k \sum_{l=1}^m \exp(-\beta \|\mathbf{x}_k - \mathbf{y}_l\|) \right\}}, \quad (14)$$

となる。したがって、 L_2 ノルムを最小化する β とは、最適化問題

$$\min_{\beta} \sum_{i=1}^n A_i \left\{ \frac{P_i}{A_i \sum_{j=1}^n P_j} - \frac{\sum_{j=1}^m \exp(-\beta \|\mathbf{x}_i - \mathbf{y}_j\|)}{\sum_{k=1}^n \left\{ A_k \sum_{l=1}^m \exp(-\beta \|\mathbf{x}_k - \mathbf{y}_l\|) \right\}} \right\}^2, \quad (15)$$

の解である。この解を β^* としよう。 β^* を求めるにあたっては、いかにも最適化問題を数値的に解く必要がある。

最適化問題の数値解法としては、ニュートン法などがよく知られている。当然、こうした方法を使うことも考えられるが、ここでは、以下に述べる非常に単純な方法を用いよう。

距離を測る単位としてメートルが用いられているとしよう。このとき、これまで行われた研究結果^{18), 19), 20), 21)} をみると、日本の都市空間においては、 β^* は 0 からせいぜい 2.0×10^3 までの範囲内に収まるように見受けられる。そのため、この範囲を十分細かく刻んで（例えば、 10^6 間隔でほぼ 2000 等分する）、それぞれを β としたときの L_2 ノルムを算出し、それらの中で最小の L_2 ノルムを与える β を、 β^* とすればよい。実際、このように最適化を行うプログラムを筆者が自作し（使用言語：C 言語）、ごく一般的なパーソナルコンピュータ（Panasonic CF-W7）で実行させても、十分に実用的な実計算時間で動作した。

$\beta = \beta^*$ のときの L_2 ノルムを L_2^* と書くことにしよう。もし、 L_2^* が十分に小さな値であると言えるならば、居住モデルがよくあてはまっていると言えよう。つまり、「施設へのアクセシビリティによって人口密度が規定されている」と解せる。また、(L_2^* が十分に小さいことに加えて) β^* が十分に大きな値であると言えるならば、施設から離れるとアクセシビリティも人口密度も急減することになる。つまり、裏を返して言えば、「施設へ近づきやすい場所で、人口が顕著に多い」となる。で

は、何と比較して L_2^* や β^* が十分に小さいであるとか、大きいであるかというのを考えてみよう。

ごく妥当な一つの考え方は、まず、「他にありえたかもしれない、今のものとは違う人口分布」を考え、次に、居住モデルのデータへのあてはめと β の推定を行い、このときにおける L_2 ノルムや β を比較基準とするというものであろう。この「他にありえたかもしれない、今のものとは違う人口分布」のことを、最初の章でも定義したように、「可能性人口分布」と呼ぶことにしよう。また、「可能性人口分布」の下で推定される距離減衰パラメータを β^{**} 、最小化される L_2 ノルムを L_2^{**} と表すことにする。

可能性人口分布がどのようなものかと問うたときに、一つのありうる答えは、空間単位 S_i の人口 P_i あるいは人口密度 p_i を、他の空間単位 S_j が持つかもしれない、逆に、 S_j の人口 P_j あるいは人口密度 p_j を、 S_i が持つかもしれない、というものであろう。つまり、空間単位と人口あるいは人口密度の組み合わせが変わったものが、可能性人口分布であるという考え方はありえよう。ただし、空間的自己相関傾向の統計的検定において仮定されるように、あらゆる空間単位と人口あるいは人口密度がどのようにでも組み合わせうるとする考え方（邦文文献では、例えば、張²⁵⁾ を参照されたい）を認めるのは不自然だろう。

もし、空間単位と人口あるいは人口密度が、まったくランダムに組み替えられたとしよう。すると、互いに近接している空間単位間で、人口の差ないし人口密度の差が著しく大きいということが、至るところで生ずる。このことは、普遍的に観察される「人は集まって住む」という事実と整合的ではなく、いかにも現実的ではない。つまり、可能性人口分布は、ある空間単位内において多くの人が居住しているのなら、それに近接している空間単位においてもやはり多くの人が居住しており、逆に、ある空間単位内にあまり多くの人が居住していないのなら、それに近接している空間単位においてもそれほど多くの人が居住していない、と考える方が自然だろう。このことを空間分析一般の用語を用いて言い換えれば、「可能性人口分布において、人口密度に関して一定の（正の）空間的自己相関が存在しているべきである」ということになる。

では、どの程度の空間的自己相関が存在すべきか考えてみよう。様々な考え方はあろうが、一つの有力な考え方は、「実際の人口分布において存在するのと同水準の空間的自己相関が存在すべき」というものだろう。一般に、空間分析においては、空間的自己相関の程度を表すのに、Moran の I 統計量が用いられる。そこで、この I 統計量を用いることにすると、「可能性人口分布における人口密度に関する I 統計量値は、実際の人口分布におけるそれと同程度であるべき」と考えられる。

人口密度に関する I 統計量値が一定となる有可能性人口分布を、解析的に扱うことはいかにも容易ではない。そこで、数値的な方法を、以下では考えていこう。

今、各空間単位に人口密度が与えられており、それに対する I 統計量を算出したとき、その値が I となることを「レベル I である」と言うことにしよう。また、手元にある人口分布データは、レベル I^* であるとしよう。レベル I^* となる有可能性人口分布をランダムに生成することを繰り返し、 L_2^{**} および β^{**} を求めることを繰り返すと、近似的に L_2^{**} および β^{**} の確率分布が得られる。ここで、 L_2^{**} 、 β^{**} それぞれの確率分布を $f(L_2^{**})$ 、 $g(\beta^{**})$ と表すことにしよう（なお、どちらも離散型確率分布である）。この $f(L_2^{**})$ および $g(\beta^{**})$ を用いることで、 L_2^* が十分に小さいか、あるいは、 β^* が十分に大きいかが、統計的に判別できる。

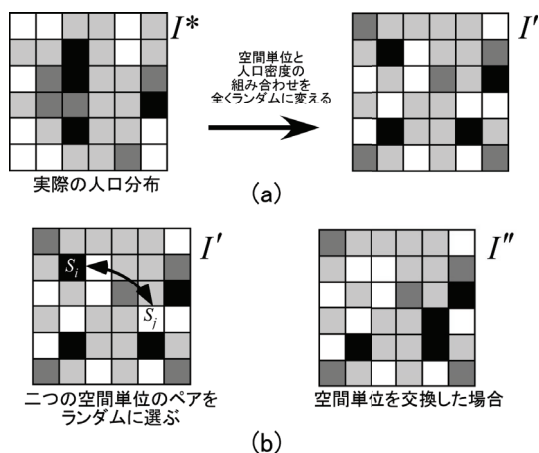


図1 有可能性人口分布をランダムに生成するプロセス

レベル I^* となる有可能性人口分布をランダムに生成するには、Goodchild²²⁾ による方法をアレンジした、以下のような方法を用いることができる。

まず、空間単位と人口密度の組み合わせを、全くランダムに変える (図1(a))。このランダムな組み合わせ変更がなされた時点で空間単位 S_i に与えられる人口密度を p_i' とする。このときにおける I 統計量を算出する。この値を I' とする。次に、二つの空間単位 S_i と S_j をランダムに選び出し、これらの中で人口密度値の交換を行うことを考える。すなわち、 S_i における人口密度を p_j' とし、 S_j における人口密度を p_i' とすることを考える。ここで、仮にこの交換を行った場合における I 統計量を算出する。この値を、 I'' と書くことにする。ここで、もし I'' の方が I' よりも I^* に近い、すなわち、

$$|I'' - I^*| < |I' - I^*|, \quad (16)$$

ならば、実際に S_i 、 S_j で人口密度の交換を行う (図1(b))。そして、この交換に伴い、これまで p_i' 、 p_j' と表してきたものを、その表記を入れ替え、それぞれ p_j' 、 p_i' と書くことにする。また、ここまで I'' と表してきた値を、今

後新たに I' と表すことにする。以下、

$$I' \approx I^*, \quad (17)$$

となるまで、空間単位の“交換”を繰り返すことにより、一つの有可能性人口分布が得られる (シミュレートされる)。

空間単位 S_i における人口密度 p_i が、空間単位間での人口密度の交換が繰り返された後に、ある空間単位 S_j に与えられたとする。このとき、 S_j の面積 A_j を、 A_i' と表すことにする。このようにすると、シミュレートされた有可能性人口分布における S 内の総人口は、

$$\sum_{i=1}^n A_i' p_i, \quad (18)$$

となる。一方、 S 内における実際の総人口は、 A_i と p_i を使うと、

$$\sum_{i=1}^n A_i p_i, \quad (19)$$

と書ける。

ここで、 $\mathbf{a} = (A_1 - A_1', A_2 - A_2', \dots, A_n - A_n')^T$ 、 $\mathbf{p} = (p_1, p_2, \dots, p_n)^T$ とする。実際の総人口と有可能性人口分布における総人口の差は、これら二つのベクトルの内積 $\langle \mathbf{a}, \mathbf{p} \rangle$ に等しい。他の考え方もあろうが、この差は 0 となるべきと考えるのが自然だろう。そこで、どのようなときに、 $\langle \mathbf{a}, \mathbf{p} \rangle = 0$ となるか考えてみよう。

\mathbf{a} と \mathbf{p} に関しては、コーシー・シュワルツの不等式

$$|\langle \mathbf{a}, \mathbf{p} \rangle| \leq \|\mathbf{a}\| \cdot \|\mathbf{p}\|, \quad (20)$$

が成り立つ。この不等式における等号は、 \mathbf{a} もしくは \mathbf{p} がゼロベクトルであり、右辺が 0 となるときに成り立つ。ベクトル \mathbf{p} がゼロベクトルではないこと (もし、そうだとすると、誰も居住していない地域を分析していることとなる) を考えると、実際の総人口と有可能性人口分布における総人口の差が必ず 0 であるためには、 \mathbf{a} がゼロベクトルでなければならない。 \mathbf{a} がゼロベクトルであるためには、任意の空間単位のペアを選んだとき、それらの面積が等しくなければならない、さらにそのためには、すべての空間単位の面積が等しくなければならない。つまり、この論文で提案する分析方法がごく自然に用いられるためには、メッシュデータのように、空間単位間で面積の差がないことが望ましい。

レベル I^* となる有可能性人口分布を繰り返し生成し (シミュレートし)、 L_2^{**} および β^{**} を求めることを繰り返すことで、近似的に $f(L_2^{**})$ 、 $g(\beta^{**})$ が求められることになる。

有可能性人口分布の生成の繰り返しにより、

$$\sum_{L_2^{**} < L_2^{(\alpha)}} f(L_2^{**}) < \alpha, \quad (21)$$

を満たす最大の $L_2^{(\alpha)}$ も近似的に求められたとしよう。すると、

$$L_2^* < L_2^{(\alpha)}, \quad (22)$$

ならば、有意水準 α %で、居住モデルは実際の人口分布データ、施設位置データに適合していると言える。つまり、この分析方法による限りでは、施設へのアクセシビリティによって人口密度が規定されていると結論付けられる。

なお、式(22)の成立の可否を判断するまでの計算手順の詳細は、以下になる。まず、有可能性人口分布の生成を $(N-1)$ 回繰り返す、各回における L_2^{**} を計算する。第 i 回目における L_2^{**} を、 $L_2^{**(i)}$ と書くことにする。 $L_2^*, L_2^{**(1)}, L_2^{**(2)}, \dots, L_2^{**(N-1)}$ を昇順にソートしたときの L_2^* の順位を R とする。すると、

$$\frac{R}{N}, \quad (23)$$

が、 $\alpha/100$ を下回るならば、式(22)が成り立つことになる。なお、式(23)で与えられる値は、ここで考えている検定に関する p 値を近似的に求めたものとなる。

同様にして、距離減衰パラメータに関する検定における p 値が(近似的に)算出されたとしよう。そして、 β^* と

$$\sum_{\beta^{(\alpha)} < \beta^{**}} g(\beta^{**}) < \alpha, \quad (24)$$

を満たす最小の $\beta^{(\alpha)}$ を比較したときに、

$$\beta^* > \beta^{(\alpha)}, \quad (25)$$

という不等式が成り立つ(と判断されることになった)としよう。このとき、有意水準 α %で、実際の人口分布データ、施設位置データから推定される距離減衰パラメータは大きいと言える。つまり、アクセシビリティは施設からの距離に応じてちどころに減衰していく、と言いうる。したがって、式(22)が成り立つと同時に、式(25)も成り立つとなれば、この論文で提案する分析方法による限りでは、施設の近くで人口が顕著に多いということになる。

3. 適用例

この章では、前章で提案した分析方法を、青森県弘前市の人口分布データ、商業施設(スーパーマーケット)の位置を記録したポイントデータへ適用した結果を示す。

分析対象領域は、弘前市全域ではなく、同市としては比較的人口密度が高く、中心部と呼ぶ領域に限定した。その領域は、図2において、太線の長方形で表されている通りである。この長方形の東西および南北方向の一辺の長さは、それぞれ4kmと5kmである。

ここで用いる人口分布データは、以下のようにして準備した。まず、長方形の分析対象領域を、一辺の長さが100mの正方形メッシュに分割した。それと並行して、ウェブサイト「政府統計の総合窓口」(URL: <http://www.e-stat.go.jp/>)より、町丁字単位で集計された平成17年度国勢調査データを入手した。このデータでは、個々の町丁字の形状はポリゴンとして記録され、それぞれのポリゴンに対し、町丁字内の人口が属性として与えられている。次に、作成したメッシュと国勢調査データを組み合わせ、面積按分法によって、(大規模都市公園内にあるなど、人口がゼロとなるべきものを除き)個々のメッシュの人口を推定した。このようにして各メッシュに対して人口が推定されたデータを、以下では、人口分布データとして用いることにする。

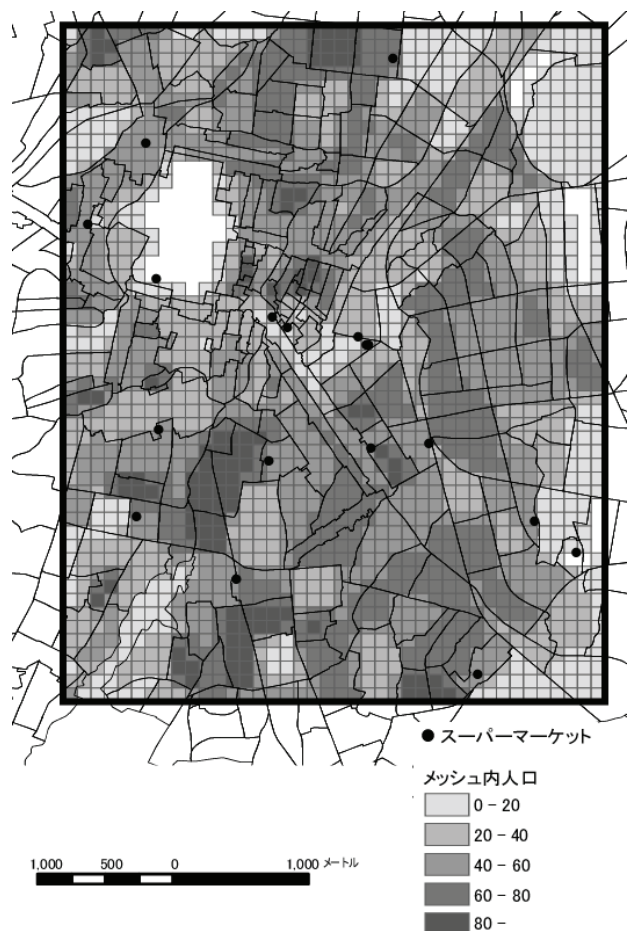


図2 分析対象領域(青森県弘前市中心部)

商業施設のポイントデータは、以下のようにして、作成した。まず、ウェブサイト「iタウンページ」(URL: <http://itp.ne.jp/>)において、「スーパーマーケット」、「弘前市」をキーワードとして、弘前市内にあるスーパーマーケットおよびそれらの住所のリストを作成した。次に、東京大学空間情報科学研究センターの提供するウェブサイト「CSV アドレスマッチングサービス」(URL: <http://newspat.csis.u-tokyo.ac.jp/geocode/>)を利用し、先

のリストからポイントデータを作成した。その後、分析対象領域内のスーパーマーケットのみを抽出した。

前章で提案した分析方法を実行に移すためには、 I 統計量を算出する必要があり、さらに、そのためには、メッシュ間の隣接関係を定める必要がある。この隣接関係を、ここでは、以下のようにして定めることにした。

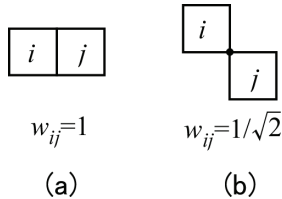


図3 メッシュ間の隣接関係の定め方

今、メッシュに対して適当に番号を振り、 i 番目のメッシュと j 番目のメッシュの隣接関係に関する量 w_{ij} を考える。もし、図3(a)に描くように、 i 番目のメッシュと j 番目のメッシュが辺を共有する場合には、 $w_{ij} = 1$ とする。図3(b)に描くように、 i 番目のメッシュと j 番目のメッシュが点を共有する場合には、 $w_{ij} = 1/\sqrt{2}$ とする。それ以外の場合には、 $w_{ij} = 0$ とする。このような w_{ij} を i 行 j 列目の要素とする行列を \mathbf{W} とする。この行列 \mathbf{W} を、行方向の和が1となるように基準化する。つまり、

$$w'_{ij} = \frac{w_{ij}}{\sum_j w_{ij}}, \quad (26)$$

なる w'_{ij} を定義したとき、この w'_{ij} を i 行 j 列目の要素とする行列を考える。この基準化によってできる行列を、 \mathbf{W}' と書くことにしよう。メッシュ間の隣接関係を \mathbf{W}' で表すものとし、また、 I 統計量の算出においても用いることにする。

行列 \mathbf{W}' によってメッシュ間の隣接関係を表すものとする、 I 統計量は、以下の式で求められる。

$$I = \frac{\sum_i \sum_j w'_{ij} (p_i - \bar{p})(p_j - \bar{p})}{\sum_i (p_i - \bar{p})}, \quad (27)$$

なお、ここで \bar{p} は、全メッシュの人口密度の平均である。また、 \mathbf{W}' は、行方向の和が1になるように基準化されているため、空間分析のテキスト (例えば、張²⁵⁾ などに見られる I 統計量の定義式よりも、式(27)は簡潔になっている。

以上の準備を踏まえ、前章で提案した分析方法を実行しよう。まず、ここで準備した人口分布データの場合、 I 統計量の値は 0.75982 と計算される。また、居住モデルを準備したデータにあてはめると、 $L_2^* = 1.276 \times 10^{-8}$ 、 $\beta^* = 3.17 \times 10^{-4}$ となった。 $N = 1000$ として、有可能性人口分布を繰り返し生成し、同様に繰り返し計算さ

れた L_2^{**} をヒストグラムにまとめたものが図4、 β^{**} をヒストグラムにまとめたものが図5である。なお、有可能性人口分布をシミュレートするにあたっては、有可能性人口分布における人口密度に関する I 統計量が、 0.75982 ± 0.00001 の範囲に収まるまで、空間単位の交換を繰り返している。先に算出された L_2^* と β^* の値に留意しながら、この図4、図5をみても、いかにも L_2^* は十分に小さく、 β^* は十分に大きいように見られる。実際、 L_2 ノルム (居住モデルの適合度) に関する検定における p 値 (式(23)によって求められる) は、0.005 にすぎず、距離減衰パラメータに関する検定における p 値も、0.006 にとどまる。以上のことから、5%あるいは1%といった通常設定される有意水準を定め、この論文で提案した分析方法を用いる限りでは、弘前市の中心部において人口分布は、スーパーマーケットへのアクセシビリティによって規定されていると言え、また、スーパーマーケットの近くで顕著に人口が多いということになる。

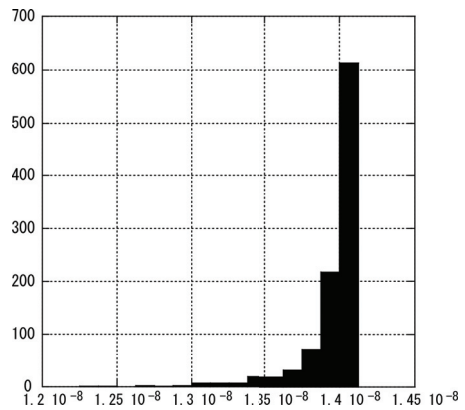


図4 有可能性人口分布における居住モデルの適合度

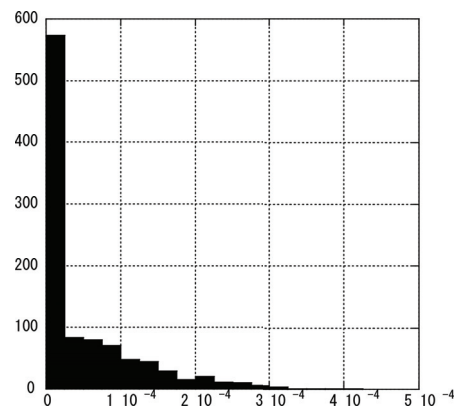


図5 有可能性人口分布における距離減衰パラメータ

4. まとめと今後の課題

この論文では、まず、人口分布が生活利便施設へのアクセシビリティによって規定されているか、そして、施設の近くで人口が顕著に多いか、を統計的に判別する方法を提案した。次に、この分析方法を、青森県弘前市の人口分布データ、スーパーマーケット位置を記録したポ

イントデータに対する結果を示した。なお、この論文で提案した方法は、一定の空間的自己相関水準を持つ空間分布をシミュレートする方法を用いることで、実行可能となっている。

この論文で提案した分析方法は、様々に拡張可能である。例えば、居住モデル以外のモデルを考えた場合にも容易に拡張可能である。また、適当な検定統計量を用いる分析方法へも容易に拡張可能である。また、人口密度と同様に、人口の増減に関しても一定水準の空間的自己相関が、一般に存在すると考えられる。つまり、人口が増加する場所の近くでは、やはり人口が同じように増加し、人口が減少する場所の近くでは、やはり人口が同じように減少する、といったことは一般的現象であろう。そこで、有可能性人口分布をシミュレートする方法を、人口指標分布全般に拡張し（この拡張も容易であろう）、人口指標分布全般を扱えるようにするということが、今後の一つの課題だろう。

現実には、生活利便施設へのアクセシビリティが一方的に人口指標分布を規定するのではなく、逆に、人口指標分布も施設立地に影響する。特に、商業施設の立地に対しては、そうした影響が大きいだろう。そこで、人口指標分布と生活利便施設立地の間の双方向的な関係を考慮しつつ、生活利便施設へのアクセシビリティが人口指標分布を規定しているかどうかを判別する分析方法を考えるということも、今後の課題として挙げられる。

最初の章で述べたように、この論文で分析方法を提案した目的は、近年の都市計画研究における関心に答えられる一つの道具立てを与えることにある。そこで、この論文における分析方法を、実際のデータを用いた実証研究で利用することも、今後の課題としたい。

参考文献

- 1) 川上 秀光 (1980) 地方都市における密度構造の変容, 都市計画論文集, 15, 73-79.
- 2) 中村 隆司 (1980) 高度経済成長期以降の地方都市における人口分布変化について—松本を対象にして, 都市計画論文集, 15, 85-90.
- 3) 佐藤 滋、戸沼 幸市 (1982) 密度を尺度とした市街地の変容過程に関する研究 (その1) 人口集中地区の変動過程の分析, 日本建築学会計画系論文集, 302, 95-106.
- 4) 大村 謙二郎、城所 哲夫、太田 守幸 (1982) 高度成長期における地方都市の市街地の人口密度構造の変化とその要因に関する研究, 都市計画論文集, 17, 7-12.
- 5) 浅見 泰司 (1982) メッシュデータを用いた DID 内外の人口分布変容の数量的分析, 都市計画論文集, 17, 85-90.
- 6) 浅見 泰司 (1983) クランピング法による幹線道路付近の人口分布の分析, 都市計画論文集, 18, 31-36.
- 7) 浅見 泰司 (1984) クランピング法による DID 及び市街化区域の境界付近の人口分布の分析, 都市計画, 131, 65-70.
- 8) 中林 一樹 (1981) 都心周辺高密度市街地の人口減少構造と人口定住化の可能性について, 都市計画論文集, 16,

253-258.

- 9) 大江 守之 (1983) 都心地域における人口減少構造—千代田区を事例として—, 都市計画論文集, 18, 259-264.
- 10) 天野 克也、松本 直司 (1991) 地方都市中心部における人口減少に関する研究, 都市計画論文集, 26, 577-582.
- 11) 白幡 武皇、樋口 秀、森村 道美 (1999) 転出者と共同住宅居住者の諸属性に着目した地方都市都心周辺部の人口減少要因分析, 都市計画論文集, 34, 661-666.
- 12) 三木 富士夫 (1983) 生活必需品を扱う小売店の分布と人口分布との関連について, 都市計画論文集, 18, 19-24.
- 13) 鈴木 勉 (1999) 移動損失基準による地域施設密度と人口密度の理論的關係に関する研究, 日本建築学会計画系論文集, 521, 183-187.
- 14) 森永 武男、有馬 隆文、萩島 哲、坂井 猛 (2000) 生活利便施設の分布から見た生活環境に関する研究, 都市計画論文集, 35, 991-996.
- 15) 海道 清信 (2001) 人口密度指標を用いた都市の生活環境評価に関する研究—交通生活及び徒歩圏の地域生活施設を中心に—, 都市計画論文集, 36, 421-426.
- 16) 橋本 晋輔、谷口 守、松中 亮治 (2009) 公共交通整備状況と地区人口密度からみた都市拡散の関連分析, 都市計画論文集, 44(1), 117-123.
- 17) 内原 英貴、吉川 徹 (2009) コンパクトシティからみた地方都市の人口社会増減の分布と生活利便性の関連分析, 日本建築学会計画系論文集, 642, 1805-1811.
- 18) 増山 篤 (2007) 商業・医療施設へのアクセシビリティと高齢者の居住パターンとの関係—千葉県浦安市を対象とした実証分析—, 都市計画論文集, 42(2), 72-79.
- 19) 増山 篤 (2008) 高齢者居住と医療機関までの距離との関係—都道府県レベル集計データを用いた統計的分析—, 日本不動産学会誌, 22(3), 114-125.
- 20) 貞広 幸雄 (1992) 都市人口分布と店舗分布の比例関係についての考察, 日本建築学会計画系論文報告集, 432, 99-104.
- 21) Okabe, A. and Sadahiro, Y. (1994) A Statistical Method for Analyzing the Spatial Relationship between the Distribution of Activity Points and the Distribution of Activity Continuously Distributed over a Region, Geographical Analysis, 26 (2), 152-167.
- 22) Goodchild, M.F. (1980) Algorithm 9: simulation of autocorrelation for aggregate data. Environment and Planning A, 12, 1073-1081.
- 23) 小西 貞則、北川 源四郎 (2004) 情報量規準, 朝倉書店 (東京) .
- 24) 河田 敬義 (1987) 情報量と統計, 統計数理, 35, 1-58.
- 25) 張 長平 (2009) 増補版 地理情報システムを用いた空間データ分析, 古今書院 (東京) .