

都市地域における任意の地点の従業員の集積と鉄道駅数の関係の分析

小西 純

Analysis of the relationship between the agglomeration of employees and the number of railway stations at any point in an urban area.

Jun Konishi

This paper uses grid square statistics to clarify the spatial distribution of "the agglomeration of employees" in the Tokyo area and estimates the relationship of it and the density of railway station. The statistical map of the estimation results was used to consider employee agglomeration.

The purposes of this paper are to clarify the relationship between the aggregated area units and the coefficient of determination of the estimation in linear regression model, and to clarify the issues concerning the aggregation method and model setting. It was confirmed that the larger the aggregated area unit, the larger the coefficient of determination. Estimating employee density need indicators for other means of transportation as independent variable as well as station density.

Keywords: 移動窓法 (moving window method), 地域メッシュ統計 (grid square statistics), 最小二乗法 (ordinary least squares), 線形回帰モデル (linear regression model)

1. はじめに

内閣府 (2012) は、我が国の製造業における市区町村ごとのデータを利用し、生産性を事業所密度で回帰することで事業所密度と生産性の関係を見ている。事業所密度が高い市区町村ほど生産性が高い結果となっており、事業所の集積は生産性と正の相関があることが述べられている。また、都道府県別の高度人材比率と労働者一人当たりの付加価値産出額の関係から、高度人材比率が高い都道府県ほど一人当たりの付加価値産出額が高いことを示している。このように、事業所や高度人材の集積は、生産性にとってプラスの効果があるとされている。

Rosenthal, S. and Strange, W. (2008) は、人的資本の外部性と、賃金と人的資本集積の関係の距離による減衰に注目し、アメリカ合衆国のデータを用いて分析を行っている。結論として、雇用の空間的中心性は賃金と正の関係があること、このような効果は中心部から距離があるほど急激に減衰することなどが挙げられている。

これらのように人的資本の集積が生産性の向上に有用であることを踏まえて、本稿では東京圏 (埼玉県、千葉県、東京都 (島嶼部を除く)、神奈川県) における「従業員の集積」について、市区町村よりも小さい地域単位で編成されている地域メッシュ統計データを利用して、その空間的な分布を明らかにするとともに、鉄道駅の密度を説明変数、従業員密度を被説明変数としてモデル化し、従業員の集積について利便性の観点から考察する。

1.1. 先行研究

都心部における従業員の集積の形状については、後藤 (1997) が事業所統計調査の 2 分の 1 地域メッシュ統計の「事業所形態別従業員数」を利用して「都市型従業員集積地区」を設定し、分析している。都市型従業員集積地区は、5000 人/km²に相当するようにメッシュ当たり 1250 人以上、及び店舗・飲食店従業員が 250 人以上という条件に加えて、メッシュ同士が連続している領域で従業員総数が 1 万人を超えるもの、と定義されている。このように従業員密度を

* 正会員 (公財) 統計情報研究開発センター (Statistical Information institute for Consulting and Analysis)
〒101-0051 東京都千代田区神田神保町 3-6 能楽書林ビル 5F E-mail: junkonishi@sinfonica.or.jp

用いて定義した全国の「都心」について、空間形状の特徴を数値化した指標等を用いて分析している。

中村・高塚（2009）は、地方に位置する政令指定都市を除く県庁所在都市を対象として、平成 12 年の国勢調査と平成 11 年の商業統計を基準とした町丁目単位のデータセットを作成し、都市内における販売額の分布が人口分布も考慮に入れた都市の空間構造によってどのように説明されるかについてモデルを定式化し、その推定を試みている。推定されたパラメータを用いて、都市内交通の改善がもたらす影響とコンパクトシティ政策を念頭においた影響に関してシミュレーション分析を行い、都市のコンパクト化（集積）が中心部の活性化に寄与することを、小売販売額の点から示している。

1.2. 目的

本稿では地域メッシュ統計と GIS を利用して任意の地域単位別に従業者数を集計し、東京圏における従業者密度の空間的な分布を地図化する。この集計した従業者密度は平滑化されるため、地域的に連担した従業者密度の分布形状を表示できる。また、これらの従業者の集積について、鉄道駅数の密度を説明変数、従業者密度を被説明変数として回帰式を作成してモデル化し、鉄道駅数、すなわち交通の利便性と、従業者密度の関係とその構造を明らかにする。最後に、各地点における集計値と回帰式による予測値との残差を地図化し、従業者密度が予測値よりも大きい地点、あるいは小さい地点を把握する。これらから、地域データを利用した回帰分析における集計地域単位と推定の決定係数の関係を明示すること、集計方法及びモデル設定に関する課題を明らかにすることを研究の目的とする。

1.3. 可変単位地区問題

地域研究では、公的統計データの都道府県別や市区町村別などの集計結果を利用した分析が多い。しかし、このような地域データを利用した分析については、集計された地域単位や分析する地域範囲のサイズで分析結果が異なる「可変単位地区問題」に注意する必要がある。地域分析においては、研究の目的に応じて分析地域単位と分析地域範囲を設定するが、地理情報を用いた分布図の作成や、各種の空間

データ解析を行う際に、分析地域単位と分析地域範囲の違いが分析結果に違いをもたらすことがある。これらの分析単位によって結果が変わる問題は空間スケール問題といわれ、可変単位地区問題（MAUP：Modifiable Areal Unit Problem）のひとつとして知られている（Openshaw（1984）、中谷（2015））。

可変単位地区問題には、分析地域単位問題や分析地域範囲問題のような空間スケール問題以外にゾーニング問題もある。ゾーニング問題とは、同じ程度のサイズの分析地域単位であっても、その分析地域単位の作り方（境界設定の仕方）によって分析結果が変わってしまうものである。本稿では回帰式の推定における集計地域単位別の決定係数やゾーニングの影響について観察する。

本稿の構成は以下の通りである。第 2 章では東京圏の市区町村別のデータを利用して従業者密度を地図化し、鉄道路線が立地し駅数が多い地域は従業者密度が高い傾向にあることを確認する。次いで市区町村別の駅密度で従業者密度を回帰する。第 3 章では地域メッシュ統計を利用して移動窓法により任意の地域単位別（バンド幅：1000m, 2000m, 3000m, 4000m, 5000m）従業者密度の分布を地図化する。それぞれの地域単位において回帰係数の推定を行い、集計サイズの違いによる決定係数の大きさを確認する。第 4 章ではバンド幅 4000m の分析地域単位のモデル式により予測値を計算し、各地点の従業者密度との残差を地図化し、駅密度と従業者密度の関係について考察する。第 5 章では今後の課題についてまとめる。

2. 既存の地域単位におけるモデル化

2.1. 利用データ

本研究では平成 28 年経済センサス活動調査の市区町村別従業者数（C～R：第 2 次産業、第 3 次産業の民営事業所）、国土数値情報の平成 28 年度鉄道データを利用し、市区町村別の分析では境界未定地域は分析から除外している。

2.2. 従業者密度と鉄道路線

図 1 は、東京圏における市区町村別従業者密度と鉄道路線を重ねたもので、この図を見ると鉄道路線

沿いの市区町村の従業者密度が高く、また、都心部から離れるほど従業者密度は低くなる傾向がある。都心部は官公庁や大企業、その他多くの事業所が立地し、鉄道網が密で交通の利便性が高いため、従業者密度が高い。駅密度以外にも従業者密度に影響を与える変数は存在すると考えられるが、本稿では分析地域単位の大きさが決定係数に与える影響を明らかにすることを目的とするため、解釈が複雑にならないように、駅密度のみを説明変数とした単回帰分析を行う。

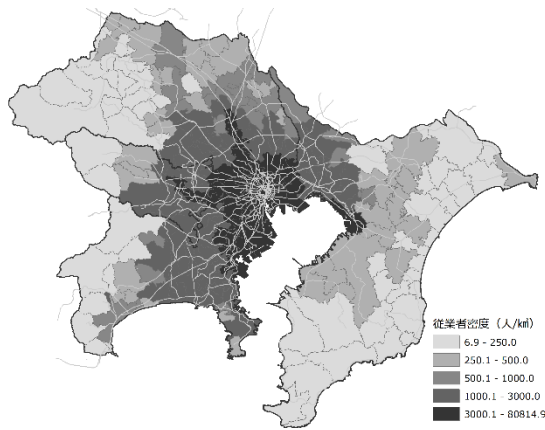


図1 市区町村別従業者密度と鉄道路線

2.3. 回帰分析

回帰分析は、説明される変数を Y で表し、これを被説明変数、説明する変数を X で表し、これを説明変数と呼ぶ。回帰分析の目的は X と Y との定量的な関係の構造（モデル）を求めることである。

ここでは、東京圏の市区町村のポリゴンに含まれる駅数と従業者数を市区町村の面積で除して駅密度と従業者密度を算出する。駅密度を説明変数 X 、従業者密度を被説明変数 Y とし、 X と Y の線形の関係进行分析する。 X と Y を関係付ける直線の傾きは、 X が1単位変化したときの Y への影響に相当し、2変数の標本データを使ってこの傾きを推計し、 X の1単位の変化が Y へ及ぼす影響を推定する。 X と Y の間の直線の傾きと切片は、最小二乗法と呼ばれる手法で推計される。

市区町村別に1 km^2 当たり1駅 (X) 減少した場合に従業者数 (Y) は1 km^2 当たり何人減少するかということを考える場合、変数の変化を定量的に把握す

ることが必要になる。その変化を β_1 とし、 $\beta_1 = 10000$ の場合、駅が1 km^2 当たり1駅減ると従業者は1 km^2 当たり10000人減少する。このような直線は以下の式で表すことができる。

$$Y = \beta_0 + \beta_1 X \quad (1)$$

ここで、 β_0 は直線の切片、 β_1 は傾きを表す。(1)式はすべての市区町村について厳密に成立する式ではなく、母集団の市区町村で平均的に成立する関係を表したものと理解される。このような線形の関係がそれぞれの市区町村で成立するには、それぞれの固有の特性（住民の特性、事業所の賃料など）といった他の要因を組み入れる必要があるが、ここではそれを「その他の要因」として表す。

今分析する市区町村数を n 、 i 番目の駅密度を X_i 、従業者密度を Y_i 、従業者密度に及ぼす「その他の要因」を u_i とすると(1)式はより一般的な表現として表される。

$$Y_i = \beta_0 + \beta_1 X_i + u_i \quad (i = 1, 2, \dots, n) \quad (2)$$

2.4. 回帰係数の推定

切片 β_0 と傾き β_1 は、母集団の回帰線における係数に当たり、切片は $X=0$ の時の母集団回帰線の値で、その点で回帰線は Y 軸と交わる。 u_i は誤差項と呼ばれる。

最小二乗推定量とは、観察されるデータに回帰線ができるだけ「近く」なるような回帰係数を選ぶもので、その「近さ」は、与えられた X に基づいて Y を予測したときの誤りを二乗した値の合計で測られる。 b_0 と b_1 をそれぞれ β_0 と β_1 の推定量とする。これらの推定量の下での回帰線は $b_0 + b_1 X_i$ で、その線に基づく Y_i の予測値も $b_0 + b_1 X_i$ となる。したがって、第 i 観測値に関する予測ミスは $Y_i - (b_0 + b_1 X_i) = Y_i - b_0 - b_1 X_i$ となりすべての n 観測値について、その予測ミスの二乗を合計したものは、次式のようになる。

$$\sum_{i=1}^n (Y_i - b_0 - b_1 X_i)^2 \quad (3)$$

この式を最小にするような切片と傾きの推定量は β_0 と β_1 の最小二乗推定量と呼ばれ、以下の式で推定できる。

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} = \frac{s_{XY}}{s_X^2} \quad (4)$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \quad (5)$$

東京圏における市区町村別の従業者密度と駅密度の関係を推定すると、推定された傾きは 12754.1、推定された切片は -766.0 となる。したがって回帰式は以下のように表せる。

$$\hat{Y} = -766.0 + 12754.1 \times \text{駅密度} (X) \quad (6)$$

2.5. 回帰式の当てはまりの指標

R^2 と回帰の標準誤差は、最小二乗推定量を使って得られた最小二乗回帰線が現実データとどの程度当てはまっているかを測る指標である。 R^2 は 0 から 1 の間の値を取り、 Y_i の分散が X_i の散らばりによってどれだけ説明されるかその割合を測る。回帰の標準誤差は、 Y_i が回帰線の予測値とどれほど離れているかを測る。

市区町村別の回帰における R^2 は 0.81 で、説明変数の駅密度は被説明変数の従業者密度の 81% を説明することを意味する。また標準誤差は 400.2 で、回帰残差の標準偏差が 400.2 であることを意味し、その単位は従業者密度である。市区町村別の従業者密度の回帰線を中心とした散らばり具合が 1 km 当たり 400 人の規模でばらついているということを意味する。

2.6. 回帰係数の検定

推定した β_1 について t 統計量による仮説検定を行うことができる。もし $\beta_1 = 0$ であれば、(6) 式の推定は意味がないということになるため、帰無仮説 $H_0: \beta_1 = 0$ を検定する。 t 統計量は以下の式で計算できる。

$$t = \frac{\text{推定量} - \text{仮説の値}}{\text{推定量の標準誤差}} = \frac{\hat{\beta}_1 - \beta_{1,0}}{SE(\hat{\beta}_1)} \quad (7)$$

(7) 式から t 統計量は 31.9 となり、帰無仮説は 0.1% 水準で棄却される。このため、駅密度は従業地密度に対して影響を与えていると統計的にみなすことができる。

2.7. 最小二乗法における仮定

最小二乗法が $\hat{\beta}_0, \hat{\beta}_1$ の適切な推定量となるための 3 つの仮定を表 1 示す。

表 1 最小二乗法における仮定

仮定 1	誤差項 u_i の X_i が与えられた下での条件付平均はゼロ
仮定 2	(X_i, Y_i) , $i = 1, \dots, n$ は独立かつ同一の分布を持つ確率変数
仮定 3	(X_i, Y_i) は有限の尖度を持つ (外れ値を持たない)

(出所) ストック, J. H. ワトソン, M. W. (2016) を参考に作成

仮定 1 は誤差項 u_i は市区町村別の従業者密度と回帰線との差であり、従業者密度を説明する「他の要因」にあたる。他の要因は回帰モデルの予想よりも高い密度 ($u_i > 0$) をもたらすこともあれば、悪い成績 ($u_i < 0$) をもたらすこともある。しかし、母集団全体で平均すれば回帰モデルの予想は正しいという状況を表す。

仮定 2 は標本データがどのように抽出されるかに関する一つの表現で、単純な無作為抽出の下で、 n 個のメンバーが母集団からランダムに選ばれ、各メンバーが選ばれる確率は同じとする。各メンバーが選ばれる可能性は皆等しく、 Y_i の分布はすべての i に対して同一なので、確率変数 $Y_1, \dots, Y_n$ は独立かつ同一の分布に従う。東京圏を母集団と考えると、市区町村別のデータは全数を分析していると考えられるが、次項で行う任意の範囲の推定では仮定 2 が成立するようにデータを無作為に抽出する。

仮定 3 はデータの外れ値に関する仮定で、最小二乗法は予想値の偏差の二乗和を最小にしているため、外れ値の影響を強く受ける。したがって外れ値を含むようなデータの分析にはあらかじめその可能性を除くなどの注意が必要である。本研究では散布図を作成して散布図で外れ値の有無を確認した。

2.8. 均一分散と不均一分散

図 2 は東京圏における市区町村別の従業者密度と駅密度の散布図と推定した回帰直線である。説明変

数が大きくなるにつれ、被説明変数の分散が大きくなっている。この対処法として頑健な標準誤差を用いることがよく行われており、頑健な標準誤差の結果を利用してt統計量を計算する。この結果、t統計量は7.8となり均一分散のみに有効なt統計量31.9よりも値は小さくなるが、p値は0.001%水準で有意となり、帰無仮説 $H_0: \beta_1 = 0$ は棄却される。

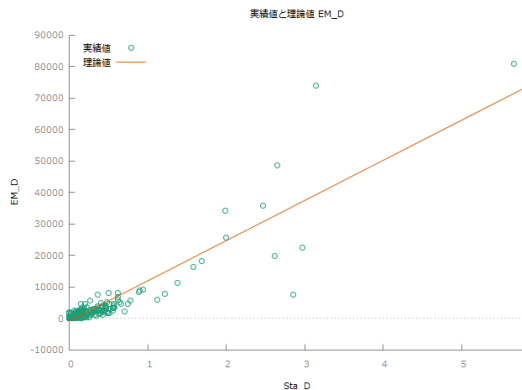


図2 市区町村別の従業員密度と駅密度の散布図

3. 任意の地域単位別の回帰推定量と決定係数

3.1. 利用データ

平成28年経済センサス活動調査の地域メッシュ統計（2分の1地域メッシュ）の従業員数（C～R：第2次産業、第3次産業の民営事業所）および国土数値情報の平成28年度鉄道データを利用した。

地域メッシュ統計データは表章地域単位であるメッシュデータのまま使用するのではなく、メッシュの図形重心点を算出し、その重心点に従業者数を収録させたものを利用した。また駅のデータは、駅を中心点座標を算出しポイントデータを作成した。

3.2. 移動窓法

作成した図形重心点データを、空間データ分析手法のひとつである「移動窓法」により任意の地域範囲に集計して利用する。空間データはある地域では大きい値を示し、他の地域では小さい値を示すなど多くの場合において場所によって異なる特徴を示すが、移動窓法を利用することにより、場所による特徴の違いを把握することができる。

移動窓法は分析地域全域の一部を覆う領域を表し、この領域（あるいは「窓」）はある位置から他の位置に移動し、窓には円形か四角形が使われることが多

い（本稿では円形を窓とした）。窓は分析範囲全体を規則正しいステップで動き、各場所で集計などの処理を実行するものとする(Lloyd,2010)。

3.3. 集計方法

従業者数、駅数については、各地域メッシュの図形重心点から半径1000m, 2000m, 3000m, 4000m, 5000m（以下バンド幅という）の円（移動窓）を作成し、その窓内にある重心点の従業者数、駅数を集計し円の面積で割って密度を算出し、その結果を重心点に収録している（図3）。

バンド幅によって集計する地域単位が異なるので、分析に当たっては移動窓の面積で従業者数を除した密度を算出した。表2に移動窓のバンド幅別の面積と東京圏における平成27年国勢調査の町丁・字等境界と市区町村の面積の平均を示す。

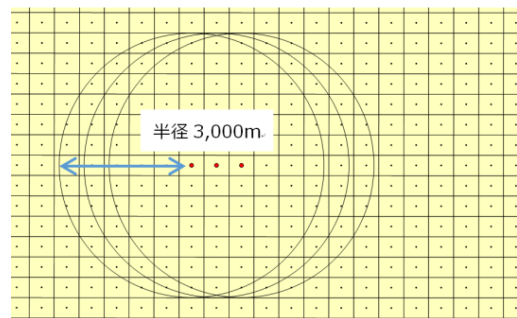


図3 メッシュの図形重心点と移動窓

表2 移動窓の面積と東京圏の集計地域区分面積の平均

バンド幅別移動窓の面積および東京圏における集計地域区分の面積の平均	面積 (km ²)
2分の1地域メッシュ(533946201)	0.26
町丁・字等(東京都島嶼部除く)	0.55
半径1000m	3.14
半径2000m	12.57
半径3000m	28.27
半径4000m	50.27
市区町村(東京都島嶼部除く)	54.34
半径5000m	78.54

3.4. バンド幅別従業員密度の分布

図4にバンド幅別従業員密度の分布を示す。バンド幅1000mの移動窓により集計した従業員密度の分布を見ると、鉄道沿線に密度が高い地域が分布している様子が目視できる。従業員密度が5000人/km²以上と高い地域は、本川越、本厚木、柏などの郊外

の市街地にも分布している。バンド幅が 3000m の移動窓により集計した従業者密度の分布を見ると、バンド幅 1000m の従業者密度の分布と比較して平滑化されており、従業者密度が 5000 人/km²以上の高い地域は東京都区部（川崎市が統合されている）、大宮駅、千葉駅、吉祥寺駅、国立駅、横浜駅など郊外のターミナル駅付近において従業者密度が高い。さらにバンド幅が 5000m の移動窓により集計した従業者密度の分布を見ると、さらに平滑化が進み従業者密度が 5000 人/km²以上と高い地域は東京都区部と川崎及び横浜が統合された地域のみとなる。

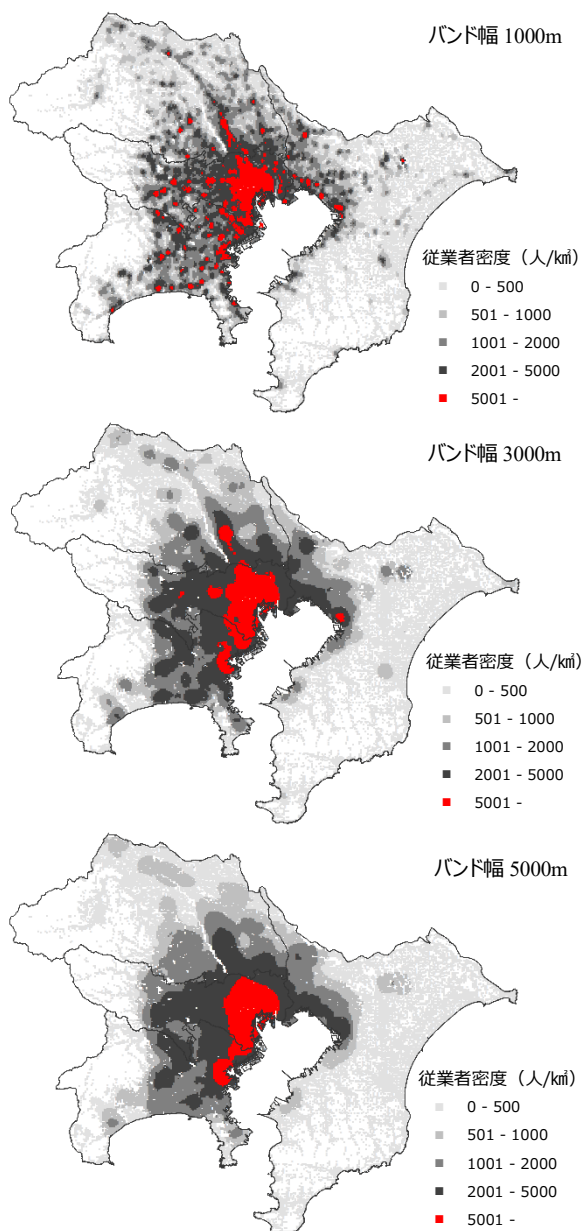


図4 バンド幅別の従業者密度の分布

3.5. ランダム抽出点

移動窓法による集計を行うと図3のように隣接する分析地域単位ではほとんどの重心点が重複するため、すべての点を使用して推定を行うと最小二乗法の仮定 2: (X_i, Y_i) の独立が満たされない。本稿では、この影響を取り除いて任意の分析地域単位別の回帰係数を推定するため、分析地域単位のサイズごとに空間的に互いに独立でランダムな点を作成し、標本を抽出した（以下、「ランダム抽出点」という）。

地域データは集計する位置によってデータ値の傾向が異なるため、多くのパターンについて東京圏における回帰係数の推定を計算することとし、ランダム抽出点はバンド幅毎に 100 組作成した。

ランダム抽出点の作成に当たっては、以下の①～③の処理を 1.5 万回繰り返して行った。

- ①GIS の乱数発生関数により任意の重心点を検索し、ランダム抽出点を示す符号を付与する。
- ②ランダム抽出点を中心としてバンド幅を半径とするバッファを作成する。
- ③バッファに重なる重心点にランダム抽出点のバッファ内であることを示す符号を付与する。

2 回目以降も①～③の処理を行うが、新しいバッファ内に既存の符号が存在しない場合にのみ新たに符号を付与する。

3.6. 回帰係数の推定

ランダム抽出点における駅密度と従業者密度から東京圏における回帰係数をバンド幅ごとに 100 組推定した。係数 β_0 、 β_1 、決定係数 R^2 の基本統計量を表3に示す。推定に当たっては頑健な標準誤差の結果を利用して t 統計量を計算している。

当てはまりの良さを示す決定係数 R^2 はバンド幅 1000m の平均は 0.604, 2000m の平均は 0.750, 3000m の平均は 0.824, 4000m の平均は 0.869, 5000m の平均は 0.902 となり、地域単位が大きくなるほど駅密度で従業者密度を回帰した際の当てはまりがよい。

係数 β_1 の推定量は t 検定の結果 100 組中 100 組の推定が 1%水準で有意である。係数 β_1 の平均をバンド幅別に見ると、8805.1 (1000m), 11776.5 (2000m), 12306.4 (3000m), 11796.5 (4000m), 11024.8 (5000m) であり、バンド幅 3000m の傾きが最も大きく、駅密

度の増加に対する従業者密度の増加度合が高くなる。

係数 β_0 の推定量は、バンド幅が小さい場合に t 検定の結果が 95%水準で有意にならないことが多く、帰無仮説 $H_0 : \beta_0 = 0$ が棄却されない。

表 3 バンド幅別最小二乗推定量および決定係数の基本統計量

	1000m		
	係数 β_0	係数 β_1	R ²
平均	-147.4	8805.1	0.604
標準誤差	10.8	59.2	0.004
標準偏差	108.5	592.1	0.039
分散	11766.8	350595.2	0.002
範囲	554.6	3313.9	0.191
最小	-407.1	6941.5	0.494
最大	147.5	10255.4	0.685
データの個数 (※)	100 (0)	100 (100)	100

() 内はp値が95%水準で有意になる数

	2000 m		
	係数 β_0	係数 β_1	R ²
平均	-443.0	11776.5	0.750
標準誤差	16.6	111.5	0.005
標準偏差	166.5	1115.0	0.045
分散	27715.8	1243289.6	0.002
範囲	635.7	4607.0	0.223
最小	-724.6	9138.9	0.616
最大	-88.9	13745.9	0.839
データの個数 (※)	100 (33)	100 (100)	100

() 内はp値が95%水準で有意になる数

	3000 m		
	係数 β_0	係数 β_1	R ²
平均	-467.7	12306.4	0.824
標準誤差	16.9	138.2	0.004
標準偏差	169.0	1382.5	0.042
分散	28547.9	1911211.6	0.002
範囲	751.0	6891.0	0.220
最小	-759.7	8099.9	0.682
最大	-8.8	14990.9	0.902
データの個数 (※)	100 (47)	100 (100)	100

() 内はp値が95%水準で有意になる数

	4000m		
	係数 β_0	係数 β_1	R ²
平均	-430.1	11796.5	0.869
標準誤差	19.6	167.5	0.005
標準偏差	196.3	1675.5	0.047
分散	38536.7	2807174.2	0.002
範囲	882.8	7996.0	0.233
最小	-796.0	7141.3	0.716
最大	86.8	15137.2	0.949
データの個数 (※)	100 (57)	100 (100)	100

() 内はp値が95%水準で有意になる数

	5000 m		
	係数 β_0	係数 β_1	R ²
平均	-349.9	11024.8	0.902
標準誤差	19.4	178.2	0.004
標準偏差	194.0	1782.0	0.040
分散	37650.7	3175483.9	0.002
範囲	805.9	7816.0	0.199
最小	-777.7	7190.3	0.765
最大	28.1	15006.3	0.964
データの個数 (※)	100 (74)	100 (100)	100

() 内はp値が95%水準で有意になる数

本稿では、東京圏の市区町村と面積の水準が最も近いバンド幅 4000m の推定量について (表 2)、100 組の平均を東京圏の母集団の推定量と考えて回帰式を作成した (式 (8))。

$$\hat{Y}_i = -430.1 + 12306.4 \times \text{駅密度} (X_i) \quad (8)$$

東京圏の市区町村別の回帰における決定係数 R² は 0.81 である。バンド幅 4000m の推定における決定係数 R² の平均は 0.87 と、バンド幅別の方が当てはまりの割合が高い傾向があるが、決定係数 R² の最小値は 0.72 と低く、ランダム抽出点の組み合わせ (抽出された地点の組み合わせ) により駅密度による従業者密度の説明力が弱くなる。

4. 推定結果の地図化

4.1. バンド幅 4000m の実測値と予測値の比較

図 5 は、バンド幅 4000m で集計された従業者密度 (実測値) と、最小二乗推定量の平均を使って得られた (8) 式にバンド幅 4000m の移動窓により計算された駅密度 X_i の下で計算される予測値 $\hat{Y}_i$ 分布である。

実測値と比較して予測値の方が従業者密度 5000 人/㎢以上の地域 (以下「高密度地域」という) が広く分布している。都心部を含めた地域は実測値よりも一回り大きく予測されており、加えて実測値では高密度ではない多摩地域 (府中市, 国立市, 日野市, 立川市), 神奈川県の藤沢市, 鎌倉市, 千葉県の市川市, 船橋市, 千葉市にも高密度地域が分布している。これらの地域は駅密度に対して従業者密度が相対的に低い地域と言える。

また予測値では駅密度が 0 の地域では従業者数に

切片の値-430 人が計算されている。従業者密度が負の値となるのが現実的ではなく、駅以外の利便性の説明変数としてインターチェンジ、バス停の密度などの他の移動手段による利便性をモデルに組みこむことや、小売業の顧客となる人口を考慮に入れるなど、モデルの定式化に課題がある。

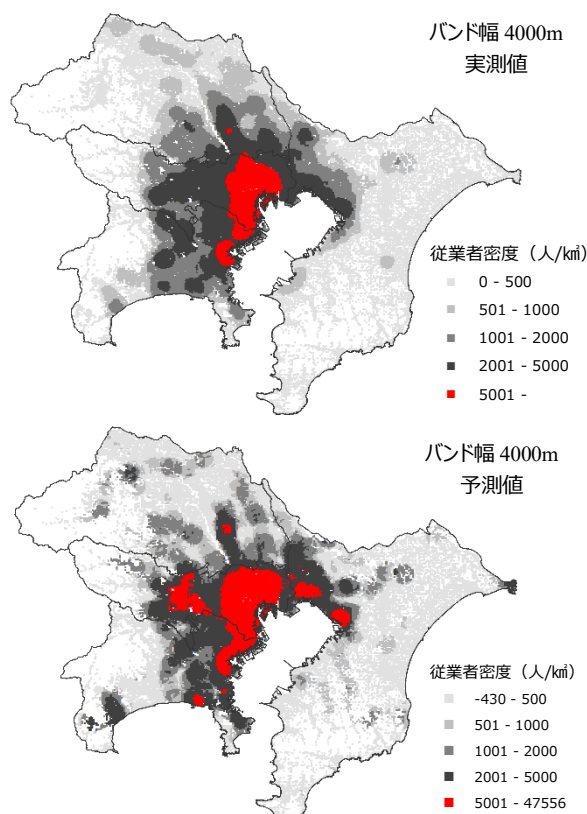


図5 従業者密度実測値と予測値の比較

4.2. 残差

バンド幅 4000m の移動窓で集計された従業者密度 Y_i と予測従業者密度 $\hat{Y}_i$ の差である残差 ($Y_i - \hat{Y}_i$) を算出し、予測値に対する割合 (%) を地図化した。東京都区部外縁部や府中、国立、立川など東京都の郊外、市川市や船橋市などでは予測値に対して実測値が小さく (青色系)、埼玉県や神奈川県西部では予測値に対して実測値が多い (赤色系)。赤色系の地域は自動車の利用が多いなど、駅密度以外の要素で従業者密度が説明される地域と考えられる。

市区町村別の結果も同様に残差率を計算して地図化すると、メッシュデータを移動窓で集計したものと同様の分布傾向が確認できるが、メッシュ別のも

のと比較して地点の情報量が少ないため、空間的に詳細な傾向を把握できない。

市区町村のような固定化されたエリアではなく、任意の地点における集計結果について残差率を地図化することにより、集計地域の大きさのみならず地点 (位置) による推定結果の違いについて表示し、より詳細に鉄道路線などの社会基盤との関係を把握できる。

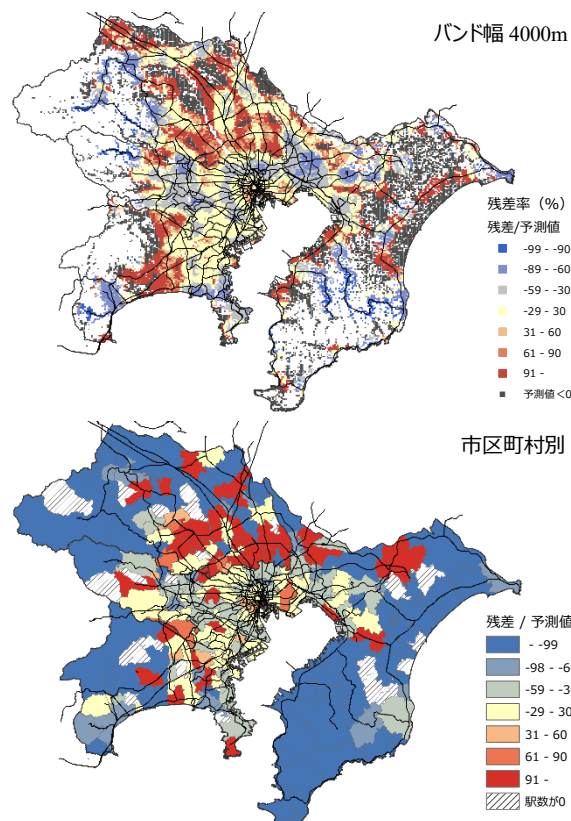


図6 残差率の分布

5. おわりに

5.1. まとめ

本稿では、市区町村別の集計結果と地域メッシュ統計を利用した任意の地域単位別の集計結果について従業者密度と駅密度の回帰式を推定した。任意の地域単位別の集計結果については、バンド幅別に決定係数を比較したところ、サイズが大きいくほど決定係数が大きくなることを確認した。

面積の水準が市区町村別と同水準にあるバンド幅 4000m の回帰係数の推定結果による予測値と実測値を比較したところ、駅密度が高いにもかかわらず従

業者密度が高くない地域が郊外にあることが確認できた。都心部への従業者の集積が著しいと考えられる。

さらに実測値と予測値について残差率をバンド幅4000mの地域単位別と市区町村別の結果で比較したところ、移動窓により任意の地点で集計した結果の方が地点の詳細な結果について把握することができ、鉄道路線などの社会基盤との関係を明らかにできることを示した。

5.2. 今後の課題

埼玉県南部や神奈川県西部などは予測値が実測値よりも小さく、駅密度のみでは回帰の度合いが小さい結果となった。駅密度に対して従業者密度が高いことを示しており、これらの地域では都心からの距離や自動車利用による移動など鉄道以外の利便性に関する説明変数を追加した重回帰分析を行うことが課題として考えられる。

今回は最も基本的な最小二乗法による回帰分析を行った。この方法は、すべての地点を含んで係数の推定を行うグローバルモデルである。定数を除く係数が地点ごとに計算されるローカルモデルである回帰分析の方法として地理的加重回帰法があり、この方法を用いて同様の分析を行うことも今後の課題として挙げられる。

参考文献

- 後藤寛 (1997) 日本における都心地域の空間形状の特性と動向, 地理学評論 70A-10 625-641
- ストック, J. H. and ワトソン, M.W. 原著, 宮尾龍蔵訳 (2016) 『入門計量経済学』, 共立出版
- 内閣府 (2012) 『平成 24 年度 年次経済財政報告ー日本経済の復興から発展的創造へー』
<https://www5.cao.go.jp/j-j/wp/wp-je12/index_pdf.html>
- 中谷友樹 (2015) . 「空間分析におけるスケール」『地理情報科学 GIS スタンダード』 浅見泰司, 矢野佳司, 貞広幸雄, 湯田ミノリ編, 古今書院.
- 中村良平・高塚創 (2009) 都市の空間構造と小売り販売額の分布ーNEG ポテンシャルモデルによる分析ー, RIETI Discussion Paper Series 09-J-022
- Lloyd, C.D.(2010) *Spatial Data Analysis : An Introduction for GIS Users*, Oxford University Press.
- Rosenthal, S. and Strange, W. (2008), *The attenuation of human capital spillovers* JOURNAL OF Urban Economics
- Stan Openshaw (1984). The Modifiable Areal Unit

Problem, 38, GeoBooks.