

小地域転出入推計の改良手法の提案と全国市区町村への適用方針検討

内田瑞生*・杉本達哉**・高森秀司**

A Proposal for an Improved Method of Estimating Social Dynamics of Small Areas and Its Application to Municipalities of Japan Based on the Census

Mizuki UCHIDA*, Tatsuya SUGIMOTO**, and Shuji TAKAMORI**

Japan is facing the challenge of ensuring the sustainability of its urban services and infrastructure as its population declines. In order to assess the sustainability, it is essential to understand the population dynamics by small area and cohort, especially the social dynamics. The purpose of this study is to estimate the number of people moving in and out by small area and cohort from aggregate data. In this study, we first examine the improvements to the estimation method proposed by Osaragi et al. (1997). Using the iterative proportional fitting (IPF), we improve the method so that the aggregated estimates are consistent with the original data. Next, to apply the improved method to all municipalities of Japan, we show the estimation procedure using the census. From the estimation results, it is found that the proposed method improves the accuracy. We also applied the proposed method to the census data and found that the estimation results are positively correlated with the actual data.

Keywords: 人口 (Population), 小地域 (Small area), 社会動態 (Social dynamics), 国勢調査 (Census), IPF 法 (Iterative proportional fitting)

1. はじめに

長期的な人口減少時代にある我が国では、地域の持続可能性を評価するうえで、人口動態の把握が重要である。人口動態は、出生・死亡に伴う「自然動態」と転出入に伴う「社会動態」に二分されるが、特に社会動態は地域差が大きく、将来人口に大きく影響することが知られている。

2014年12月に制定された「まち・ひと・しごと創生総合戦略」では、東京圏への一極集中の是正を目的として地方移住が推進されている。さらに、昨今の自然災害の激甚化や COVID-19 の流行に伴い、居住地選択のあり方に関心が集まっている。

また、同年7月に発表された「国土のグランドデザイン2050」では、生活・社会サービスなどの都市機能の集約化と各都市機能へのアクセシビリティの両立を目指す「コンパクト+ネットワーク」の実現が掲げられており、地域間連携を念頭に置いた都市機能の配置検討が求められている。

今後、都市機能の配置検討を行ううえでは、小地域などの詳細な地域区分で人口分布や人口推移を分析することが有効である。そのため、社会動態についても、小地域単位でのデータ整備が求められる。さらに、社会動態は、ライフステージと密接な関係にあるため、コーホート（男女・年齢階級）別のデータであることが望ましい。

近年、小地域（町丁目）の人口関連データの公開が進んでいるが、小地域・男女別の転入者数の合計値などの公表に留まっており、小地域・コーホートごとの特性に留意した社会動態分析には限界がある。

そこで、本研究では、オープンデータを活用した「小地域・男女・年齢階級別」の転出者数および転入者数の推計手法を提案する。まず、一部都市が公表する実測値による妥当性検証を踏まえて、大佛ら（1997）の提案した従来手法の改善提案を行う。その後、提案手法を全国に適用することを可能とするため、オープンデータを用いた推計方針を提案する。

* 正会員 八千代エンジニアリング株式会社 技術開発研究所 (Yachiyo Engineering Co., Ltd.)
〒111-8648 東京都台東区浅草橋 5-20-8 E-mail : mz-uchida@yachiyo-eng.co.jp

** 非会員 八千代エンジニアリング株式会社 技術開発研究所 (Yachiyo Engineering Co., Ltd.)

2. オープンデータ・先行研究

2.1. 人口関連データのオープンデータ

「政府統計の総合窓口 (e-Stat)」では、国勢調査などの人口データが多数公開されており、市区町村以上の地域区分のデータの多くは、データベースからのダウンロードが可能となっている。

一方で、小地域人口関連データは、e-Stat「地図で見る統計 (統計 GIS)」からダウンロード可能な国勢調査の男女・年齢階級別人口や e-Stat の国勢調査小地域集計からダウンロード可能な小地域・男女別の転入者数データなどに限られ、小地域・男女・年齢階級別の社会動態データは公開されていない。

また、一部の市区町村は、独自に、住民基本台帳に基づく小地域別の転出者数および転入者数を公開している。しかし、公開されているデータのほとんどが、年齢階級を属性に含まないデータであるため、年齢階級を含む詳細な社会動態データを公開している市区町村は、千葉県柏市などのごく一部の市区町村に限られる。

2.2. 先行研究

先行研究として、小地域・男女・年齢階級別の転出者数および転入者数を、各属性に対する集計値から推計した大佛ら (1997) が挙げられる。大佛ら (1997) の手法では、仮想データに対して、高い推計精度が得られることが確認されているが、実データに対する妥当性検証は行われていない。また、年齢階級別の移動人口の集計値に関する整合性が担保されていないという課題がある¹⁾。

また、分割表の集計値 (以下、周辺分布と呼ぶ) から集計化前のデータを推計する手法として、Deming ら (1940) が提案した Iterative Proportional Fitting (以下、IPF 法) が広く知られている。IPF 法の適用例としては、国勢調査データに適用した浅見ら (1998)、地域間人口移動の推計を行った斎藤ら (1990) などが挙げられる。IPF 法では、制約条件が単一の分割表で閉じていることが求められるが、本研究で扱う社会動態は、転出入それぞれに制約条件が課されるため、直接的な利用は困難である。

3. 提案手法

3.1. 従来手法の改善提案

大佛ら (1997) の提案手法 (以下、従来手法と呼ぶ) では、年齢階級別の移動人口および定住人口の周辺分布に対する整合性が担保されておらず¹⁾、4.1 の実測値を対象とした妥当性検証でも周辺分布に対する不整合が確認された。

そこで、本研究では、全ての周辺分布に対して整合的な手法を提案する。従来手法と同様に、小地域・男女・年齢階級別の転出者数および転入者数を、各属性の集計値からそれぞれ推計する。ここでは、2時点間の人口移動を対象とし、1 時点目の人口と生残率を考慮した 2 時点目の封鎖人口および 2 時点目の人口を、それぞれ、移動前人口および移動後人口と定義する。さらに、移動前後の人口を、それぞれ、2 時点間に移動しなかった定住人口および転出または転入を行った移動人口に二分する。

ここで、男女それぞれに対して、移動の有無 i (0: 定住, 1: 移動)・小地域 j ・年齢階級 k 別の移動前人口 $l_{ijk} \geq 0$ および移動後人口 $m_{ijk} \geq 0$ を定める。このとき、定住人口の定義より、任意の j, k に対して、 $l_{0jk} = m_{0jk}$ が成り立つ。また、周辺分布を、 $x_{+jk} = \sum_i x_{ijk}$ などと表す。すなわち、 l_{ijk} および m_{ijk} の周辺分布はそれぞれ表 1 に示されたデータと対応する。

大佛ら (1997) が指摘するように、転出者数および転入者数は、定住人口から算出可能である。ここで、定住人口は、移動前後のいずれの人口も超えてはならないため、実質的な上限制約を与えるのは移動前後人口のうち小さい方の人口である。

表 1 移動前後の周辺分布と対応データ

周辺分布		データ
移動前	l_{+jk}	小地域・年齢階級別移動前人口
	l_{i+k}	年齢階級別定住人口 ($i = 0$) 年齢階級別転出者数 ($i = 1$)
	l_{ij+}	小地域別定住人口 ($i = 0$) 小地域別転出者数 ($i = 1$)
移動後	m_{+jk}	小地域・年齢階級別移動後人口
	m_{i+k}	年齢階級別定住人口 ($i = 0$) 年齢階級別転入者数 ($i = 1$)
	m_{ij+}	小地域別定住人口 ($i = 0$) 小地域別転入者数 ($i = 1$)

そこで、定住人口に関する制約条件を表現するため、新たに、 $n_{ijk} \geq 0$ を式 (1) および式 (2) が成り立つように定める。

$$n_{0jk} = l_{0jk} = m_{0jk} \quad (1)$$

$$n_{+jk} = \min(l_{+jk}, m_{+jk}) \quad (2)$$

以下に、 n_{ijk} の周辺分布 n_{i+k} および n_{ij+} が、表 1 に示した周辺分布より導出可能であることを示す。

まず、 n_{0jk} の周辺分布 n_{0+k} および n_{0j+} は、式 (1) および表 1 に示した周辺分布より、それぞれ、式 (3) および式 (4) のように表される。

$$n_{0+k} = l_{0+k} = m_{0+k} \quad (3)$$

$$n_{0j+} = l_{0j+} = m_{0j+} \quad (4)$$

さらに、 n_{1jk} の周辺分布 n_{1+k} および n_{1j+} は、表 1 に示した周辺分布、式 (3) および式 (4) より、式 (5) および式 (6) のように表される。

$$n_{1+k} = \sum_j n_{+jk} - n_{0+k} \quad (5)$$

$$n_{1j+} = \sum_k n_{+jk} - n_{0j+} \quad (6)$$

n_{ijk} の周辺分布 n_{+jk} 、 n_{i+k} および n_{ij+} を制約条件として推定した \hat{n}_{ijk} は、転出入いずれの制約条件も満たす。

以上より、転出入それぞれの制約条件は、 n_{ijk} に関する単一の分割表に対する制約条件に帰着することが示される。そのため、既存の分割表推定手法である IPF 法の適用が可能となる。

本研究では、3次元の分割表に対して、IPF 法を適用する。IPF 法の手順を以下に示す。

【Step1】 初期値 $n_{ijk}^{(0)}$ (時点 $t = 0$) を与える。

【Step2】 式 (7)、式 (8) および式 (9) より時点 t の推定値 $n_{ijk}^{(t)}$ から時点 $t + 1$ の推定値 $n_{ijk}^{(t+1)}$ を算出する。

$$n'_{ijk} = n_{ijk}^{(t)} \frac{n_{+jk}}{n_{+jk}^{(t)}} \quad (7)$$

$$n''_{ijk} = n'_{ijk} \frac{n_{i+k}}{n_{i+k}'} \quad (8)$$

$$n_{ijk}^{(t+1)} = n''_{ijk} \frac{n_{ij+}}{n_{ij+}''} \quad (9)$$

【Step3】 $\max |n_{ijk}^{(t+1)} - n_{ijk}^{(t)}|$ が予め設定した収束基準値 (本研究では、 10^{-3} を用いた) を満たさない場合には、**【Step2】**を繰り返す。収束基準値を満たした時点 s の推定値 $n_{ijk}^{(s)}$ を最終的な推定値 \hat{n}_{ijk} とする。

また、本研究では、式 (10) に示すように、人口規模に市区町村全体での定住人口または移動人口の割合を掛けた値を、初期値として用いた。

$$n_{ijk}^{(0)} = n_{+jk} \frac{n_{i+k}}{n_{+k}} \quad (10)$$

本研究では、IPF 法の推定に、Barthélemy ら (2018) の開発した R パッケージである `mipfp` を利用した。

以上の手法により、従来手法の課題であった、一部周辺分布に対する不整合への対処が可能となった。

3.2. オープンデータを用いた推計手法の提案

3.1 は、表 1 に示すデータの利用を前提とした手法であるが、多くの市区町村では、一部データしか公表されていないという課題がある。したがって、提案手法を全国の市区町村に適用するためには、オープンデータを用いたデータ生成が必要となる。

そこで、オープンデータを用いた推計方針を以下に、示す。ここでは、全国データが公開されている国勢調査および市区町村別生命表を使用する。

3.1 の提案手法は、移動前後の人口 l_{+jk} 、 m_{+jk} および n_{ijk} の周辺分布が与えられれば推定可能である。本研究では、国勢調査小地域集計で転入者数のみが公開されていることや転入が居住地選択行動をより直接的に表すことを踏まえて、入手が容易な移動前後の人口および転入者数を用いた推計方針をとる。

ここでは、ある t 年の国勢調査と、その5年後の $t + 5$ 年の国勢調査および市区町村別生命表を用いて、 t 年から $t + 5$ 年にかけての小地域・男女・年齢階級別の転出者数および転入者数を推定するものとする。ただし、年齢階級として、0歳から75歳以上までの5歳階級を用いる。

小地域・男女・年齢階級別の移動前人口は、 t 年における人口に $t + 5$ 年における男女・年齢階級別の生残率を掛けて算出する²⁾。ただし、生残率は、市区

町村・男女・年齢階級別に一律に与える。また、 $t+5$ 年時点の0~4歳人口は推計対象から除外する。ここで、転入者数は、国勢調査の5年前の常住地の集計区分のうち、移動人口に分類される国内（自市区町村内・自市内他区・県内他市区町村・他県）、国外の和で与える。ただし、5年前の常住市区町村「不明」（5年前の常住地が他市区町村だが市区町村名が不明の者）は、通常、移動人口とみなされないため、ここでは転入者数に含めない。さらに、市町村合併や小地域区分の変更が行われた地域では、合併前人口を重みづけとする生残率の加重平均の処理やジオメトリ面積比を用いた人口配分などが必要となる。

使用するオープンデータの2000年以降のデータ整備年および対応する周辺分布を表2に示す。

オープンデータを用いた推計方針を以下に示す。ただし、推計は男女それぞれに対して行われる。

【Step1】 t 年の小地域・年齢階級別人口に $t+5$ 年の市区町村別生命表より算出した生残率を掛けて、 l_{+jk} を算出する。 $t+5$ 年の小地域・年齢階級別人口を m_{+jk} とする。 n_{+jk} を式(2)で算出する。

【Step2】 年齢階級別転入者数 m_{1+k} は式(11)を満たす必要がある。そのため、年齢階級別転入者数データが式(11)の変域に含まれない場合には最も近い端点に一致させる³⁾。その後、年齢階級別定住人口 $n_{0+k} = l_{0+k} = m_{0+k}$ および n_{1+k} を、それぞれ式(12)および式(13)により算出する。

$$\sum_j m_{+jk} - \sum_j n_{+jk} \leq m_{1+k} \leq \sum_j m_{+jk} \quad (11)$$

$$n_{0+k} = \sum_j m_{+jk} - m_{1+k} \quad (12)$$

$$n_{1+k} = \sum_j n_{+jk} - n_{0+k} \quad (13)$$

【Step3】 小地域別転入者数 m_{1j+} は、式(14)および式(15)を満たす必要がある³⁾。そのため、小地域別転入者数データに対して補正を行い、 m_{1j+} を算出する（後ほど詳述する）。その後、小地域別定住人口 $n_{0j+} = l_{0j+} = m_{0j+}$ および n_{1j+} を、それぞれ式(16)および式(17)により算出する。

$$\sum_k m_{+jk} - \sum_k n_{+jk} \leq m_{1j+} \leq \sum_k m_{+jk} \quad (14)$$

$$\sum_j m_{1j+} = \sum_k m_{1+k} \quad (15)$$

$$n_{0j+} = \sum_k m_{+jk} - m_{1j+} \quad (16)$$

$$n_{1j+} = \sum_k n_{+jk} - n_{0j+} \quad (17)$$

【Step4】 IPF法により周辺分布 n_{+jk} 、 n_{i+k} および n_{ij+} から小地域・年齢階級別の定住人口 $\hat{n}_{0jk} = \hat{l}_{0jk} = \hat{m}_{0jk}$ を推定する。その後、小地域・年齢階級別の推計転出者数 \hat{l}_{1jk} および推計転入者数 \hat{m}_{1jk} を、それぞれ式(18)および式(19)より算出する。

$$\hat{l}_{1jk} = l_{+jk} - \hat{n}_{0jk} \quad (18)$$

$$\hat{m}_{1jk} = m_{+jk} - \hat{n}_{0jk} \quad (19)$$

ここで、**【Step3】**における小地域別転入者数データの補正方法について解説する。簡便な方法として、小地域別転入者数データに、共通の補正係数を掛ける方法が考えられるが、元データに0が存在する場合に、制約条件が満たされない可能性がある。そこで、内分比を用いた補正方法を以下に提案する。

補正前の小地域別転入者数データを m'_{1j+} と表す。ここで、 m'_{1j+} は、座標 $\sum_k m_{+jk}$ と座標 $\sum_k m_{+jk} - \sum_k n_{+jk}$ を結ぶ線分を $p'_j: (1-p'_j)$ に内分する点と捉えられる。ただし、 p'_j は、式(20)で表される。

ここで、 p'_j に共通の補正項を足した値を p_j とする。ただし、補正值が、変域 $0 \leq p_j \leq 1$ に含まれない場合には、 p_j を最も近い端点、すなわち0または1に一致させ、 $\sum_k n_{+jk} = 0$ の場合は、 $p_j = 0$ とする。

表2 使用データの概要

使用データ	整備年	周辺分布
地図で見る統計 国勢調査 小地域・男女・年齢階級 (5歳階級)別人口	2000年~ 2015年 (5年毎)	l_{+jk} m_{+jk}
市区町村別生命表		
国勢調査 5年前の常住地・ 男女・年齢階級(5歳階級)別 人口—市区町村	2000年 2010年 2015年	m_{1+k}
国勢調査 小地域集計 5年前の常住地・ 男女別人口—町丁・字等	2000年 2010年 2015年	m_{1j+}

最後に、 $p_j(1-p_j)$ を新たな内分比とする m_{1j+} を式(21)より算出する。

以上の算出結果が式(15)を満たすように適切な補正項を定める⁴⁾。本研究では、Rのuniroot関数を用いて求根を行った。

$$p'_j = \frac{\sum_k m_{+jk} - m'_{1j+}}{\sum_k n_{+jk}} \quad (20)$$

$$m_{1j+} = \sum_k m_{+jk} - p_j \sum_k n_{+jk} \quad (21)$$

4. 妥当性検証方法

4.1. 大佛ら(1997)の改善手法の妥当性検証方法

3.1の提案手法の妥当性検証におけるテストデータとして、千葉県柏市が住民基本台帳に基づいて公開している、2008年度から2019年度までの12年分の小地域(大字・町丁)・男女・年齢別人口、転出者数および転入者数を用いる。

住民基本台帳の人口移動データは、住民基本台帳への登録時点のデータであり、集計方法が国勢調査とは異なることを踏まえて、双方のデータ比較は相関係数に基づいて行うものとする。また、住民基本台帳と国勢調査の地域区分は必ずしも一致しない。本研究では、国勢調査を用いた推計結果に対するテストデータとして利用するため、今回は、国勢調査の小地域と同名の小地域を対象とする(295地域のうち約92.5%にあたる273地域が対象となった)。

柏市は、千葉県北西部に位置し、2015年度時点で人口は約41万人である。柏市の2015年度の人口密度および鉄道路線の状況を図1に示す。次に、柏市の移動人口推移および2015年度の年齢階級別移動人口を、それぞれ、図2および図3に示す。図2より、転出者数は約2,500人程度の振れ幅で推移している一方で、転入者は、2012年と2019年で5,000人以上の差が存在することがわかる。また、図3より、年齢階級によって移動人口が大きく異なり、柏市では、移動人口が20歳から34歳までに集中しており、25~29歳の移動人口が最も多いことがわかる。

妥当性検証として、まず、テストデータから表1に示した周辺分布データを作成し、従来手法と3.1

の提案手法のそれぞれで集計前のデータを推計する。その後、それぞれの推計結果の制約条件の満足状況および推計精度の比較を行う。

4.2. 全国適用方針の妥当性検証方法

次に、3.2の推計方針に従い、2010年国勢調査、2015年国勢調査および2015年市区町村別生命表を用いて、2010年から2015年にかけての柏市の小地域・男女・年齢階級別の転出者数および転入者数を推計する。その後、テストデータにおける2010年度から2014年度までの5年分の転出者数および転入者数の累計値との比較を行う。

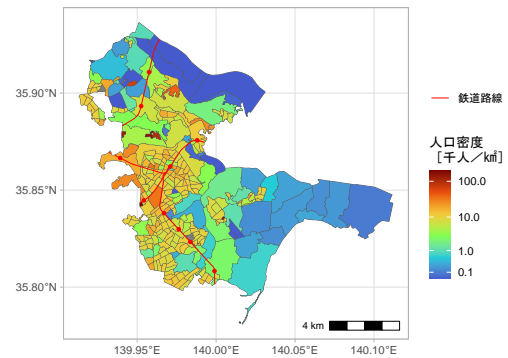


図1 柏市の人口密度と鉄道路線状況(2015年度)

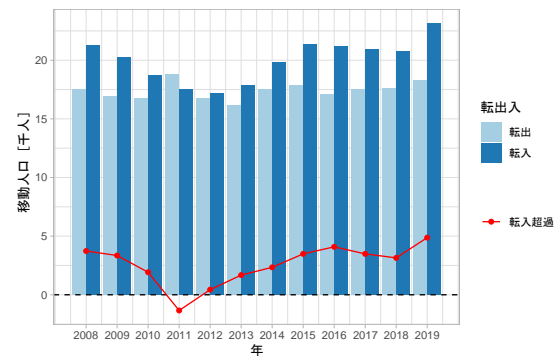


図2 柏市の移動人口推移

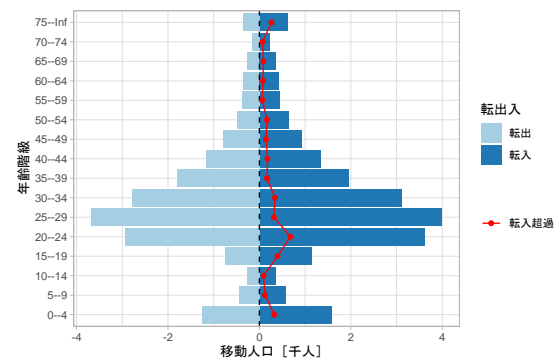


図3 柏市の年齢階級別移動人口(2015年度)

また、移動人口は、人口規模と比例関係を有することが予想される。そのため、移動人口を人口規模との比例関係のみで説明する回帰モデルよりも、今回、推計した人口移動率などを加味したモデルのほうが、高い説明力を有することが望ましい。そこで、式 (22) および式 (23) に示す、2 種類のポアソン回帰モデルを設定する。ただし、 λ_{ijk} は、 l_{ijk} または m_{ijk} のいずれかに対応する。

ここで、式 (22) は、単に移動人口が定住人口に比例すると仮定したモデルである。一方で、式 (23) は、式 (22) のモデルに推計人口移動率の項を加えたモデルである。妥当性検証として、これらのモデルを対象にカイ二乗検定に基づく分散分析を行い、説明力に有意な差が生じるかを検証する。

ただし、本研究では、人口規模を表す定住人口を、テストデータにおける 2010 年度から 2014 年度までの平均定住人口で与えた。

$$\ln \lambda_{ijk} = \ln \lambda_{0jk} + \beta_0 \quad (22)$$

$$\ln \lambda_{ijk} = \ln \lambda_{0jk} + \beta_0 + \beta_1 \frac{\hat{\lambda}_{1jk}}{\lambda_{0jk}} \quad (23)$$

5. 妥当性検証結果

5.1. 大佛ら (1997) の改善手法の妥当性検証結果

従来手法と 3.1 の提案手法の推計結果を図 4 に示す。図 4 より、従来手法と提案手法の双方で、推計移動人口と実績移動人口間に高い正の相関がみられるが、提案手法の相関係数が、従来手法をわずかに上回っていることがわかる。

従来手法における年齢階級別の転出者数 l_{1+k} および転入者数 m_{1+k} に関する制約条件の満足状況を表 3 に示す。ここでは、男女ともに同様の傾向がみられるためデータを集計化している。また、転出入それぞれの誤差は等しくなることに注意が必要である。

表 3 より、各年齢階級において平均で数十人程度の誤差が生じており、周辺分布に対する整合性が保たれていないことが確認できる。さらに、提案手法では、全ての時点および年齢階級で制約条件が満たされることが確認された (全ての時点および年齢階級に対して誤差区間が $\pm 10^{-3}$ 以内)。

また、従来手法と 3.1 の提案手法の RMSE は、それぞれ、1.40 [人]、1.38 [人] であり、12 時点中 8 時点で提案手法の推計精度が従来手法を上回った。また、従来手法と 3.1 の提案手法の MAE は、それぞれ、0.628 [人]、0.623 [人] であり、12 時点中 10 時点で提案手法の推計精度が従来手法を上回った。

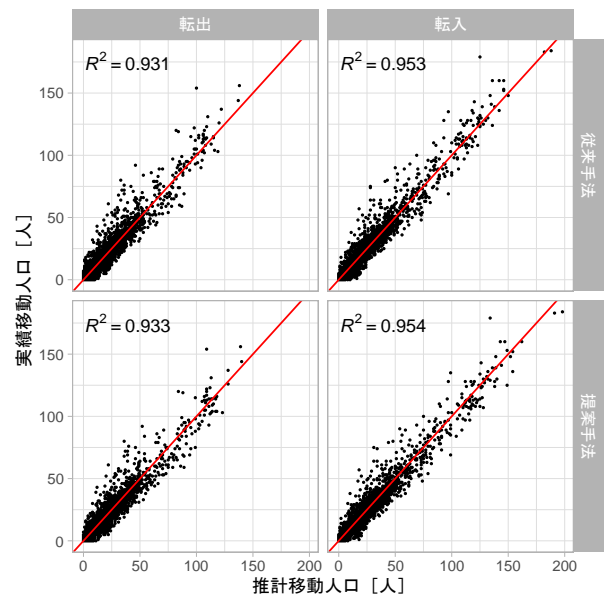


図 4 従来手法と提案手法における小地域・男女・年齢階級別移動人口の推計結果

表 3 従来手法の年齢階級別移動人口の制約条件満足状況

年齢階級	平均誤差 [人]	誤差区間 (最小～最大) [人]
5～9 歳	17.7	9.2 ～ 23.3
10～14 歳	20.5	13.6 ～ 27.6
15～19 歳	-0.8	-15.1 ～ 10.5
20～24 歳	-72.0	-110.1 ～ -43.3
25～29 歳	-50.0	-79.0 ～ -21.3
30～34 歳	-21.9	-39.4 ～ 5.7
35～39 歳	-14.4	-32.0 ～ 3.4
40～44 歳	-1.5	-16.9 ～ 10.9
45～49 歳	5.4	-10.6 ～ 22.2
50～54 歳	13.9	6.5 ～ 21.3
55～59 歳	15.5	5.0 ～ 26.4
60～64 歳	15.2	6.2 ～ 22.7
65～69 歳	18.3	9.3 ～ 28.8
70～74 歳	18.1	10.6 ～ 27.1
75 歳以上	16.7	4.4 ～ 30.0

以上より、3.1 の提案手法は、従来手法の制約条件上の課題を解決することが確認できた。また、わずかではあるが、推計精度の向上にも寄与しうることが確かめられた。

5.2. 全国適用方針の妥当性検証結果

最も移動人口の多い 25～29 歳における小地域別の推計移動人口と実績移動人口の関係を、図 5 に示す。図 5 より、転出入いずれも高い正の相関がみられることがわかる。

年齢階級ごとの推計移動人口と実績移動人口の相関係数およびテストデータより算出した定住人口と実績移動人口の相関係数を図 6 に示す。

図 6 より、転出入ともに、大半の年齢階級において、相関係数が 0.75 を超える高い正の相関がみられることがわかる。また、一部の年齢階級では、定住人口の有する相関係数よりも高い正の相関がみられることから、人口規模の説明力よりも高い説明力を有している可能性が示唆される。

また、転出入のいずれにおいても高齢層の相関係数が低い傾向がみられ、特に転出の 75 歳以上においてその傾向が顕著であった。これは、高齢層の移動人口が少ないことに起因する可能性がある。

最後に、式 (23) の回帰モデルのパラメータ推定結果および式 (22) のモデルとの分散分析（カイ二乗検定）の p 値の結果を表 4 および表 5 に示す。

表 4 および表 5 より、65 歳以上の転出者数および

75 歳以上の転入者数を除く年齢階級では、人口規模の影響を打ち消してもなお、推計人口移動率が移動人口に正の効果をもつことが確認でき、分散分析でも有意な差がみられた。

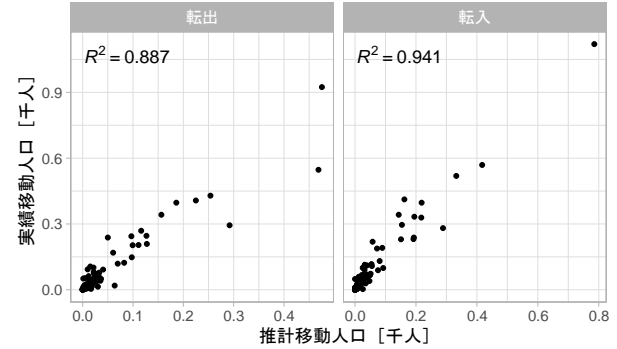


図 5 25～29 歳の移動人口の推計値と実績値の関係

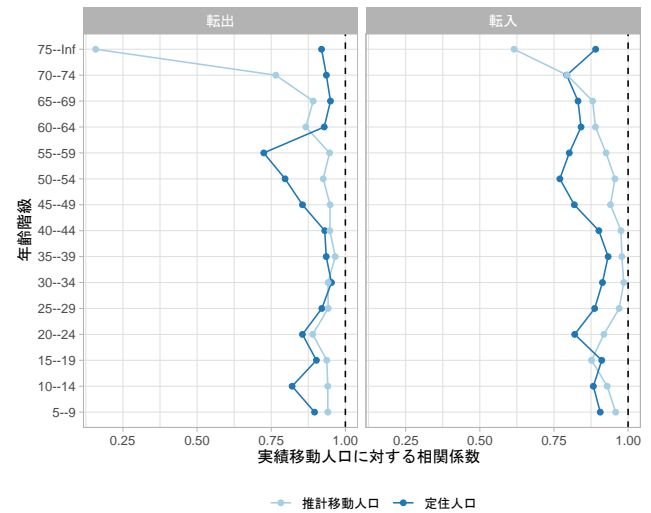


図 6 推計移動人口および定住人口の実績移動人口に対する年齢階級別の相関係数

表 4 転出に関するポアソン回帰モデル推定結果

年齢階級	β_0 : 切片	β_1 : 推計移動率	分散分析
5～9 歳	-1.66***	1.08***	***
10～14 歳	-2.49***	1.61***	***
15～19 歳	-2.66***	2.30***	***
20～24 歳	-1.54***	2.65***	***
25～29 歳	-0.47***	0.57***	***
30～34 歳	-0.59***	0.43***	***
35～39 歳	-1.27***	0.94***	***
40～44 歳	-1.80***	1.40***	***
45～49 歳	-2.25***	2.01***	***
50～54 歳	-2.53***	2.45***	***
55～59 歳	-2.68***	1.59***	***
60～64 歳	-2.82***	0.98***	***
65～69 歳	-3.02***	0.43	
70～74 歳	-3.22***	-0.16	
75 歳以上	-3.11***	-0.44	

*** : $p < 0.001$, ** : $p < 0.01$, * : $p < 0.05$

表 5 転入に関するポアソン回帰モデル推定結果

年齢階級	β_0 : 切片	β_1 : 推計移動率	分散分析
5～9 歳	-1.79***	1.59***	***
10～14 歳	-2.54***	2.35***	***
15～19 歳	-2.41***	1.38***	***
20～24 歳	-0.92***	0.54***	***
25～29 歳	-0.60***	0.81***	***
30～34 歳	-0.88***	0.93***	***
35～39 歳	-1.10***	0.69***	***
40～44 歳	-1.76***	1.42***	***
45～49 歳	-2.18***	1.97***	***
50～54 歳	-2.28***	1.73***	***
55～59 歳	-2.48***	1.78***	***
60～64 歳	-2.69***	1.33***	***
65～69 歳	-2.95***	1.66***	***
70～74 歳	-3.16***	1.98***	***
75 歳以上	-2.69***	0.01	

*** : $p < 0.001$, ** : $p < 0.01$, * : $p < 0.05$

6. おわりに

本研究では、大佛ら (1997) の小地域・男女・年齢階級別の移動人口の推計手法に対して、IPF 法を用いた改善手法を提案した。さらに、国勢調査を用いて、提案手法を全国に適用する方針を示した。

実績値に基づく妥当性検証により、提案手法は、全ての周辺分布に関する制約条件を満たすと同時に、推計精度の向上に寄与することがわかった。また、国勢調査を用いた推計移動人口が、実績値に対して高い正の相関を有することが確かめられた。

本研究は、データ生成に重きを置いているが、今後は、小地域・コーホートを考慮した居住地選択モデルの構築など、推計データの活用方針について検討を行う予定である。また、データ生成上の課題として、他の分割表推定手法との推計精度比較を行っていないことや推計値の整数への変換について検討されていないことが挙げられる。

脚注

1) 大佛ら (1997) では、転出入それぞれの従う確率分布については明記されていない。ここでは、転出入をいずれも二項分布に従うと仮定することで、従来手法に一致する結果が得られることを示す。本文中の表記に従えば、転出入はそれぞれ式 (24) および式 (25) で表される二項分布に従う。ただし、 p_k および q_k は、それぞれ、市区町村全体での転出率および転入率であり、式 (26) で表される。いま、転出入間の独立性を仮定すれば、同時確率は、式 (27) となる。さらに、従来手法と同様にラグランジュの未定乗数法およびスターリングの近似を用いれば、1 次の必要条件が式 (28) のように求められ、これは、従来手法の結果と一致する。以上より、年齢階級別の移動人口に関する周辺分布 l_{1+k} および m_{1+k} は、あくまで転出率 p_k または転入率 q_k として確率に反映されているに過ぎず、制約条件としては扱われていないことがわかる。

$$P_k(n_{0jk}) = \frac{l_{+jk}!}{n_{0jk}!(l_{+jk} - n_{0jk})!} (1 - p_k)^{n_{0jk}} p_k^{l_{+jk} - n_{0jk}} \quad (24)$$

$$Q_k(n_{0jk}) = \frac{m_{+jk}!}{n_{0jk}!(m_{+jk} - n_{0jk})!} (1 - q_k)^{n_{0jk}} q_k^{m_{+jk} - n_{0jk}} \quad (25)$$

$$p_k = \frac{l_{1+k}}{l_{++k}}, \quad q_k = \frac{m_{1+k}}{m_{++k}} \quad (26)$$

$$\prod_k [P_k(n_{0jk}) Q_k(n_{0jk})] \quad (27)$$

$$-2 \ln n_{0jk} + \ln(l_{+jk} - n_{0jk}) + \ln(1 - p_k) - \ln p_k + \ln(m_{+jk} - n_{0jk}) + \ln(1 - q_k) - \ln q_k = 0 \quad (28)$$

2) x 歳階級の生残率は、70 歳未満の 5 年後生残率 $S_{x<70}$ および 70 歳以上の 5 年後生残率 $S_{x \geq 70}$ に分けて、式 (29) より算出した。ただし、 ${}_yL_x$ は、 $x \sim x + y$ 歳の定常人口を表す。

$$S_{x<70} = \frac{{}_5L_{x+5}}{{}_5L_x}, \quad S_{x \geq 70} = \frac{{}_\infty L_{75}}{{}_\infty L_{70}} \quad (29)$$

3) 補正は、移動人口および定住人口がいずれも正でなくてはならないという条件を満たさない一部のデータに対してのみ行われる。【Step2】では、 m_{+jk} および n_{+jk} を所与として、 m_{1+k} および n_{1+k} を算出する。いま、式 (30) が成り立つため、 $n_{1+k} \geq 0$ および式 (30) より式 (11) が得られる、また、同様に、式 (14) が得られる。4.2 の妥当性検証では、75 歳以上男性の転入者数のみが式 (11) を満たさなかったため、元データの 1,623 人 (転入率約 8.6%) を約 1,844 人 (転入率約 9.8%) に補正した。また、男女・小地域の 546 通りの組合せのうち 22 通り (約 4.0%) で、式 (14) が満たされなかった。このうち、変域から最も大きく外れていたのは、豊四季台 3 丁目の女性データであり、元データが 700 人 (転入率約 49.7%) であったのに対して、約 817 人 (転入率約 58.0%) の下限が与えられた。ただし、ここでの転入率は、移動後人口に占める転入者数の割合を表す。

$$n_{0+k} = \sum_j m_{+jk} - m_{1+k} = \sum_j n_{+jk} - n_{1+k} \geq 0 \quad (30)$$

4) 異なるデータ間の不整合が大きくなければ、内分比の補正項の絶対値は小さくなる。4.2 の妥当性検証では、内分比の補正項は、男性で約 2.44%、女性で約 2.36% であった。

参考文献

- 浅見泰司・木戸浩司（1998）国勢調査住宅関連統計の IPF 法による度数分布表推計の精度：東京大都市圏を例として。「日本建築学会計画系論文集」, **63** (514), 185-189.
- 大佛俊泰・前島一夫（1997）小地域における転出入人口の推計とその空間分布特性。「GIS-理論と応用」, **5** (1), 1-9.
- 斎藤参郎・長井剛一郎・本村裕之・石橋健一（1990）不完全データのもとでの年齢別男女別地域間人口移動テーブルの推定法。「都市計画論文集」, **25**, 349-354.
- Barthélemy, J., & Suesse, T. (2018). mipfp: An R Package for Multidimensional Array Fitting and Simulating Multivariate Bernoulli Distributions. *Journal of Statistical Software*, 86(Code Snippet 2), 1 - 20.
- Deming, W., & Stephan, F. (1940). On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known. *The Annals of Mathematical Statistics*, **11** (4), 427-444.