

開発などによる突発的人口変動を考慮した小地域別将来人口推計

7. 都市計画—10. 調査・評価と計画支援

正会員 内田瑞生*¹

会員外 木暮洋介*¹

小地域 将来人口推計 平滑化
ベイズ推定 異常検知 コーホート要因法

1. はじめに

現在、我が国は、長期的な人口減少過程にあり、今後、高度経済成長期前後に整備された社会資本の老朽化が急速に進むことが懸念されている。こうした人口減少社会において、社会資本整備の効率性を高めていくためには、町丁目などの詳細な地域区分（以下、小地域と呼ぶ）でのインフラ需要把握が必要であり、今後、小地域別の将来人口推計の重要性が高まると考えられる。

我が国では、国立社会保障・人口問題研究所（社人研）による将来推計人口が 5 年ごとに作成・公表されているが、小地域別の将来推計人口は作成・公表されていない。一方で、小地域を対象とした将来人口推計ツール・事例は、国土技術政策総合研究所（国総研）¹⁾による「将来人口・世帯予測ツール V2（H27 国調対応版）」、国土政策局²⁾による「メッシュ別将来人口推計」や井上³⁾による「全国小地域別将来人口推計システム」など、すでに複数存在する。

これらの先行事例では、コーホート変化率法・コーホート要因法のいずれかが用いられているが、こうした小地域別の将来人口推計の課題として、人口の小さい地域において、移動率・コーホート変化率などが不安定になるという少数問題（small number problem）⁴⁾が知られている。また、古藤⁵⁾は、開発・建設といった不連続的人口変動（以下、突発的人口変動と呼ぶ）がみられることを推計上の課題に挙げている。これらは人口の過大推計を引き起こすため、小地域別の将来人口推計で特に重要な課題となる。

こうした課題に対処するため、国総研¹⁾の推計事例では、対象となる小地域の純移動率が 2 以上の場合に、市区町村別の平均純移動率で置き換えるという処置がなされており、国土政策局²⁾の推計事例では、市区町村別の純移動率のみが用いられている。しかし、純移動率に上限を定める処置については、適切な基準が存在せず、暫定的な対応といえる。一方で、平均純移動率を用いる処置についても、小地域の地域特性が反映されないという課題が生じる。

少数問題に対する代表的なアプローチとして、ベイズ推定に基づく平滑化が知られており、空間疫学において死亡率の推定などに活用されてきた^{たとえば 6), 7)}。一方で、

井上³⁾は、ベイズ推定に基づく平滑化手法では、異常値に対する平滑化の効果が弱く、将来人口推計に適用することは困難であると指摘し、人口ポテンシャル⁸⁾に基づく平滑化手法を新たに提案した。

ベイズ推定に基づく平滑化手法の平滑化効果の弱さは、これらの手法が、突発的人口変動の影響の除去を目的としたものでないことに起因すると考えられる。通常、ベイズ推定に基づく平滑化では、少数問題の解決に主眼が置かれ、特定期間における観測ノイズのみが除去される。一方で、将来人口推計では、将来の移動率やコーホート変化率の推定が必要となるため、データの観測期間における観測ノイズだけでなく、当該期間での突発的人口変動に対しても、別途、平滑化を行う必要があると考えられる。

そこで本研究では、突発的人口変動を考慮したベイズ推定に基づく平滑化手法を提案し、従来の将来人口推計手法との精度比較を行う。具体的には、混合分布に基づく異常検知をベイズ推定に基づく平滑化手法に応用することで、少数問題の解決に加えて、突発的人口変動の影響の除去を可能とする手法を提案した。

2. 将来人口推計の方法

2.1. コーホート変化率法・コーホート要因法

ここでは、最も一般的な将来人口推計方法のコーホート変化率法とコーホート要因法について述べる。

コーホート変化率法は、コーホート要因法の簡便法に位置付けられるが、コーホート要因法と同程度の推計精度を有することが知られている⁹⁾。一方で、コーホート変化率法では、人口変動が死亡および転出入のいずれによるかが区別されないため、死亡数や転入超過数の推計が必要な場合には、コーホート要因法が用いられる。時点 t から時点 $t+n$ にかけての地域 i ・性別 s ・年齢階級 a （ n 歳階級）のコーホート変化率 ${}^s c_a^t$ は、式(1)で定義される。ただし、年齢階級 a を、各階級の最低年齢で表す。

$${}^s c_a^t = \frac{{}^s p_a^{t+n}}{{}^s p_a^t} \quad (1)$$

${}^s p_a^t$: 時点 t ・地域 i ・性別 s ・年齢階級 a の人口

コーホート要因法は、コーホート変化率を生残率と移動率に分解することで、自然増減（出生・死亡）と社会増減（転入超過）をそれぞれ推計可能な方法である。

コーホート要因法では、地域*i*・性別*s*・年齢階級*a*の人口が時点*t*から時点*t+n*にかけて生き残る割合である生残率 ${}^sS_a^t$ が、式(2)で定義される。生残率は、社人研が公表する市区町村別生命表の定常人口などから算出可能である。また、コーホート要因法では、通常、コーホート変化率から生残率を差し引いた純移動率が設定される。

$${}^sS_a^t = \frac{{}^sP_a^{t+n} - {}^sD_a^t}{{}^sP_a^t} \quad (2)$$

${}^sD_a^t$: 時点*t*から時点*t+n*にかけての死亡数

本研究では、コーホート要因法に基づき将来人口推計を行うこととする。ただし、ベイズ推定に基づく平滑化では、観測される事象数がポアソン分布に従うと仮定することが一般的であり⁴⁾、これらの発生率は正である必要がある。そのため、負になりうる純移動率に対してベイズ推定に基づく平滑化を行うことは困難である。そこで、本研究では、負になりうる純移動率ではなく、式(3)で定義される移動率 ${}^s m_a^t$ を用いる。したがって、式(4)のように、コーホート変化率 ${}^s c_a^t$ は生残率 ${}^s S_a^t$ と移動率 ${}^s m_a^t$ の積として表される。

$${}^s m_a^t = \frac{{}^s P_a^{t+n}}{{}^s P_a^t - {}^s D_a^t} \quad (3)$$

$${}^s c_a^t = {}^s S_a^t {}^s m_a^t \quad (4)$$

また、子ども女性比 ${}^s b^t$ は、式(5)で表される。ただし、本研究では、子ども女性比を男女別に算出することにより出生性比を考慮した。

$${}^s b^t = \frac{{}^s P_0^{t+n}}{{}^s F^{t+n}} \quad (5)$$

${}^s F^{t+n}$: 時点*t+n*・地域*i*の出生可能年齢女性人口

以上より、時点*t*から時点*t+n*にかけての人口変動が、式(6)で表される。小地域別の将来人口推計では、移動率 ${}^s m_a^t$ や子ども女性比 ${}^s b^t$ が過大となった場合に人口の過大推計が生じる可能性がある。そこで、次節以降では、これらを平滑化する手法について述べる。

$$\begin{cases} {}^s P_a^{t+n} = {}^s S_a^t {}^s m_a^t {}^s P_a^t \\ {}^s P_0^{t+n} = {}^s b^t {}^s F^{t+n} \end{cases} \quad (6)$$

2.2. 将来人口推計における従来の平滑化手法

国総研¹⁾や国土政策局²⁾による先行事例では、推計対象の小地域を含む地域区分（市区町村や都道府県など）で集計した移動率を用いることで平滑化を行っている。こうした方法は、地域化手法（regionalization methods）の一種に位置付けられる⁴⁾（以下、平均化手法と呼ぶ）。

井上³⁾は、ベイズ推定に基づく平滑化手法の平滑化の効果の弱さを指摘したうえで、人口ポテンシャル⁸⁾に基づく平滑化手法を提案した（以下、井上手法と呼ぶ）。井上手

法では、小地域*i*の移動率（または子ども女性比） $\hat{\theta}_i$ が、式(7)により算出される。ただし、以下では、ある時点の人口に生残率を掛けた人口を封鎖人口と呼び、その次の期の人口を次期人口と呼ぶこととする。また、小地域*i*以外の添え字は省略する。

$$\hat{\theta}_i = \begin{cases} \frac{N' + \sqrt{n_i' N'}}{N}, & \text{if } n_i = 0 \\ \frac{\sqrt{n_i} n_i' + \sqrt{n_i} N'}{\sqrt{n_i + \sqrt{N}} n_i + \sqrt{n_i + \sqrt{N}} N'}, & \text{if } n_i \neq 0 \end{cases} \quad (7)$$

n_i' : 小地域*i*の次期人口（出生数）

n_i : 小地域*i*の封鎖人口（出生可能年齢女性人口）

最後に、ベイズ推定に基づく従来の平滑化手法を取り上げる。本研究では、Anselinら⁴⁾と同様に、次期人口（または出生数）がポアソン分布に従い、移動率（または子ども女性比）がガンマ分布に従うと仮定する。これは、式(8)のように表せる。

$$\begin{aligned} n_i' | \theta_i &\sim \text{Poisson}(\theta_i n_i) \\ \theta_i &\sim \text{Gamma}(\alpha, \beta) \end{aligned} \quad (8)$$

$\theta_i > 0$: 移動率（子ども女性比）

$\alpha > 0, \beta > 0$: パラメータ

式(8)において、次期人口（または出生数）が n_i' のとき、移動率（または子ども女性比） θ_i の事後分布は、ベイズの定理より式(9)となる⁴⁾。そのため、 θ_i の事後分布の期待値より、移動率（または子ども女性比）が、式(10)のように平滑化される。ただし、ポアソン分布の母数を定める n_i は正でなくてはならないため、 $n_i = 0$ の場合には、事前分布の期待値により平滑化を行うものとする。

$$\theta_i | n_i' \sim \text{Gamma}(n_i' + \alpha, n_i + \beta) \quad (9)$$

$$\hat{\theta}_i = \begin{cases} \frac{n_i' + \alpha}{n_i + \beta}, & \text{if } n_i > 0 \\ \frac{\alpha}{\beta}, & \text{if } n_i = 0 \end{cases} \quad (10)$$

式(10)に示した、ベイズ推定に基づく従来の平滑化手法により、少数問題に対処することが可能であると考えられる。一方で、将来人口推計に特有の課題である突発的人口変動の影響の除去は困難であると考えられる。

2.3. 突発的人口変動を考慮した平滑化手法の提案

本節では、突発的人口変動の影響を除去可能な平滑化手法を提案する。

全国などを対象に将来人口推計を行う場合、移動率・子ども女性比に突発的人口変動による異常値が含まれる可能性が非常に高いといえる。このように、データに異常値が含まれることが明確な場合には、混合分布を用いたクラスタリングによる異常検知が有効である¹⁰⁾。

本研究では、移動率（または子ども女性比） θ_i が、正常値および異常値、それぞれに対応する2つの確率分布を混合した混合ガンマ（gamma mixture）分布に従うと仮定

し、式(11)のモデルを提案する。混合分布による異常検知では、確率分布ごとに算出される負担率（合計が 1 となる）に基づき、正常値か異常値かが判断される。具体的には、あらかじめ閾値を定め、正常値の確率分布の負担率が閾値を上回る場合には、正常値であると判断する。

$$n_i' | \theta_i \sim \text{Poisson}(\theta_i n_i)$$

$$\theta_i \sim \text{GammaMixture}(\phi, \alpha_{In}, \beta_{In}, \alpha_{Out}, \beta_{Out})$$

0.5 < ϕ ≤ 1: 正常値の確率分布の混合係数 (11)

$\alpha_{In} > 0, \beta_{In} > 0$: 正常値の確率分布のパラメータ
 $\alpha_{Out} > 0, \beta_{Out} > 0$: 異常値の確率分布のパラメータ

本研究では、正常値の確率分布の負担率が 0.5 以上の場合に、正常値であると判断した。式(11)より、移動率（または子ども女性比）が、式(12)のように平滑化される。

$$\hat{\theta}_i = \begin{cases} \frac{n_i' + \alpha_{In}}{n_i + \beta_{In}}, & \text{if } n_i' \text{が正常値} \\ \frac{\alpha_{In}}{\beta_{In}}, & \text{if } n_i' \text{が異常値または } n_i = 0 \end{cases} \quad (12)$$

以下では、ベイズ推定に基づく平滑化手法のうち、前節の式(10)の従来手法を単一分布手法、式(12)の提案手法を混合分布手法と呼んで区別する。本研究では、ガンマ分布とポアソン分布の混合分布が負の二項分布になることを利用し、最尤法により、確率分布のパラメータ推定を行った。

3. 将来人口推計の精度検証

3.1. 将来人口推計の実施概要

本研究では、2010 年・2015 年の国勢調査の小地域・男女・5 歳階級別人口および社人研による市区町村別生命表より、全国の 2045 年までの小地域・男女・5 歳階級別の将来人口推計を行う。

ただし、小地域の地域区分については、2020 年国勢調査の実績値による精度検証が可能となるよう、2020 年国勢調査の小地域に合わせる処理を行った。また、年齢階級については、市区町村・男女・5 歳階級（0～95 歳以上）別人口に基づく按分処理により、0 歳から 75 歳以上までの 5 歳階級を 0 歳から 95 歳以上までの 5 歳階級に変換した。さらに、女性の出生可能年齢を、社人研による地域別将来人口推計（2018 年推計）と同様に、15 歳から 49 歳までと設定した。

本研究では、単一分布手法、混合分布手法、平均化手法および井上手法の 4 種類の手法で将来人口推計を実施した。ただし、単一分布手法および混合分布手法では、移動率に関するパラメータ推定を男女・年齢階級別に行い、子ども女性比に関するパラメータ推定を男女別に行った。また、平均化手法および井上手法では、市区町村および都道府県を集計区分として、それぞれ推計を行った。

将来人口推計の結果、混合分布手法では、移動率の約 2.8 %、子ども女性比の約 0.23 %が異常値と判定された。

3.2. 将来人口推計の精度検証結果

まず、長期推計における人口の過大推計の発生状況を確認するため、社人研による地域別将来人口推計（2018 年推計）の 2045 年時点の将来推計人口（市区町村・男女・5 歳階級別）と各手法の推計結果を比較した。

社人研による 2045 年将来推計人口との比較の結果を図-1 に示す。図-1 より、ベイズ推定に基づく従来手法である単一分布手法では、井上³⁾が指摘した通り平滑化の効果が弱く、人口の過大推計が生じていることが確認できる。また、市区町村を集計区分とする平均化手法・井上手法でも一部、人口の過大推計が生じた。一方で、混合分布手法や、都道府県を集計区分とする平均化手法・井上手法では、人口の過大推計は生じなかった。

次に、単一分布手法における人口の過大推計が開発による突発的人口変動に起因することを確かめるため、国土地理院の公開する空中写真を用いて、開発状況の確認を行った。ここでは、2015 年人口に対する 2045 年人口の比を小地域別に算出し、上位 20 小地域を対象に、2010 年前後から 2015 年前後にかけての空中写真を取得した。空中写真が取得できた 16 小地域について確認したところ、いずれの小地域においても 2010 年前後から 2015 年前後にかけて開発が行われたことが明らかとなった。

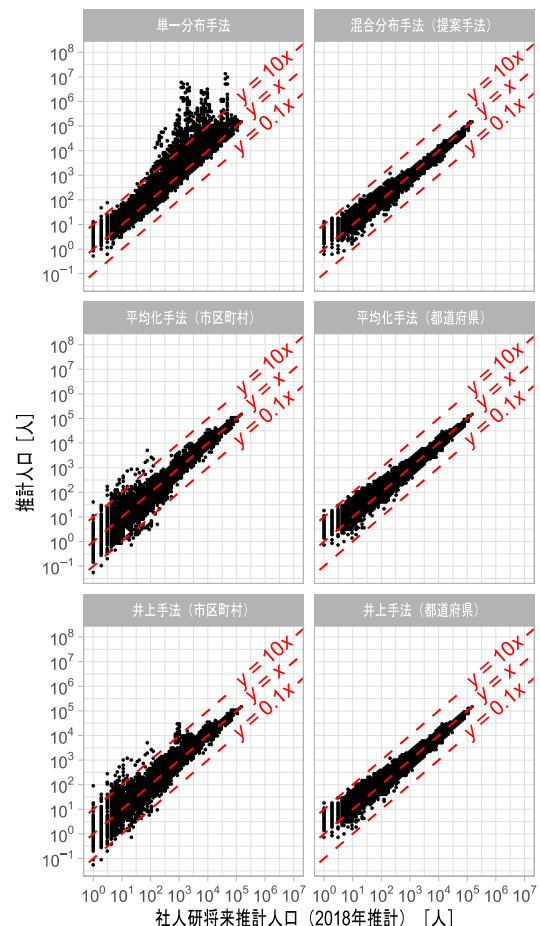


図-1 社人研将来人口の 2045 年人口との比較
 （市区町村・男女・5 歳階級別）

単一分布手法の人口比上位 20 地域のうち、上位 5 地域における、単一分布手法および混合分布手法の人口比の結果を、表-1 に示す。また、参考として、上位 20 地域のうち、下位 5 地域（第 16 位から第 20 位）の結果を表-2 に示す。表-1 および表-2 より、単一分布手法では、2015 年から 2045 年までの 30 年間で、人口が約 700 倍から約 17,000 倍になるといった人口の過大推計が生じていることが確認できる。一方で、混合分布手法では、これらの小地域の人口比は、約 1.2 から約 1.6 となり、開発に伴う突発的人口変動の影響が除去されていることがわかる。例として、人口比第 1 位の福岡市西区西都一丁目および第 20 位の名古屋市名東区高針荒田の空中写真を、それぞれ、図-2 および図-3 に示す。

最後に、2020 年国勢調査の小地域・男女・5 歳階級別人口による精度検証を行った。ここでは、小地域・男女・5 歳階級別人口に対する平均絶対誤差（MAE）および平均絶対パーセント誤差（MAPE）を都道府県別に算出した。ただし、MAPE の算出では、真値 0 のデータを無視した。精度検証の結果を図-4 に示す。図-4 より、混合分布手法の MAE および MAPE は、単一分布手法・平均化手法・井上手法の MAE および MAPE よりも小さくなった。

以上より、混合分布手法が、突発的人口変動に伴う人口の過大推計の解消に有効であることに加えて、従来手法よりも高い推計精度を有することが確認された。

4. おわりに

本研究では、開発などによる突発的人口変動が移動率などに与える影響を除去するため、混合分布による異常検知を活用した平滑化手法を提案し、将来人口推計の精度検証を行った。検証の結果、従来手法の課題であった長期推計における人口の過大推計が生じないことや、従来手法よりも推計精度が高いことが確認された。

本研究では、男女・5 歳階級ごとに移動率などの異常検知を行っているため、学生寮のように異なる年齢階級間で転出入が相殺される社会増減が異常と判定されてしまう。こうした問題への対処は今後の課題としたい。

参考文献

- 1) 国土技術政策総合研究所：将来人口・世帯予測ツール V2 (H27 国調対応版)，G 空間情報センター，2021
- 2) 国土政策局：メッシュ別将来人口推計，国土交通省
- 3) 井上孝：「全国小地域別将来人口推計システム」正規版の公開について，E-Journal GEO, Vol.13, No.1, pp.87-100, 2018
- 4) Anselin, L., Lozano, N., and Koschinsky, J.: Rate Transformations and Smoothing. Spatial Analysis Laboratory, Department of Geography, University of Illinois Urbana-Champaign, 2006
- 5) 古藤浩：小地区短期間多地域データからの地区成分解析 山形市町丁目人口データを対象とした分析と人口予測，都市計画論文集, Vol.43, No.3, pp.61-66, 2008
- 6) Tango, T.: Empirical Bayes Estimation for Mortality Indices, Japanese Journal of Applied Statistics, Vol.17, No.2, pp.81-96, 1988
- 7) 府川哲夫，清水時彦：小地域生命表のベイジアン・アプローチ，人口学研究, Vol.13, pp.37-49, 1990
- 8) Inoue, T.: A new method of estimating small area demographics

using population potential, Working Paper Series, Institute of Economic Research, Aoyama Gakuin University, 2014-3, pp.1-16, 2014

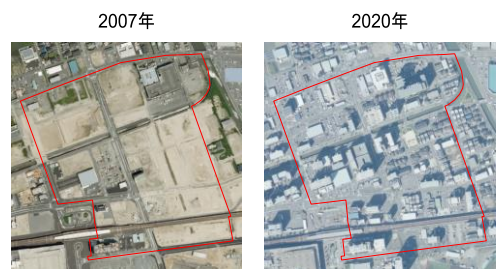
- 9) 日本人口学会：人口大事典，培風館，2002
- 10) 井手剛：入門 機械学習による異常検知 ―R による実践ガイド―，コロナ社，2015

表-1 単一分布手法人口比第 1～5 位の結果比較

順位	都道府県名	市区町村名	小地域名	2015年に対する2045年人口比	
				単一分布手法	混合分布手法
1	福岡県	福岡市西区	西都一丁目	16,688	1.22
2	大阪府	摂津市	南千里丘	15,663	1.42
3	埼玉県	越谷市	レイクタウン八丁目	9,355	1.56
4	福岡県	福岡市西区	徳永北	7,454	1.21
5	愛知県	長久手市	西原山	5,396	1.43

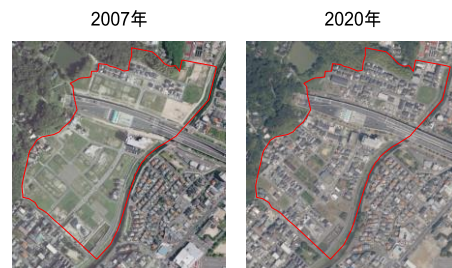
表-2 単一分布手法人口比第 16～20 位の結果比較

順位	都道府県名	市区町村名	小地域名	2015年に対する2045年人口比	
				単一分布手法	混合分布手法
16	茨城県	つくば市	みどりの中央	1,020	1.44
17	大阪府	八尾市	龍華町一丁目	1,006	1.19
18	福岡県	福岡市西区	今宿西	921	1.50
19	茨城県	つくば市	学園南三丁目	809	1.64
20	愛知県	名古屋市名東区	高針荒田	743	1.34



国土地理院撮影の空中写真に小地域境界を追加して掲載

図-2 福岡市西区西都一丁目の空中写真



国土地理院撮影の空中写真に小地域境界を追加して掲載

図-3 名古屋市名東区高針荒田の空中写真

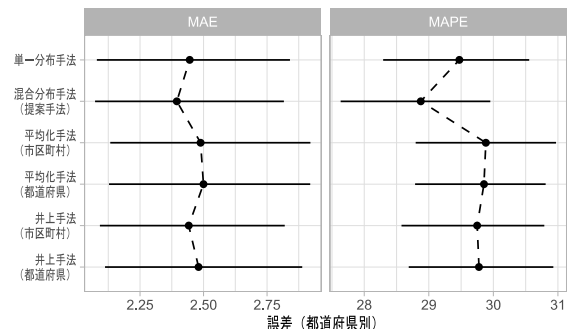


図-4 2020 年国勢調査人口による精度検証結果