

オープンデータを用いた初期世帯マイクロデータの生成に関するシステム構築

名城大学 学生会員 ○平野 巧真

名城大学 正会員 鈴木 温

1. はじめに

我が国では、少子高齢化、人口減少が進行し、市街地の空き家の増加や都市機能低下が懸念されている。そのような問題の解決策を考える上で、小規模な地域ごとに人口変化を予測することはきわめて重要である。その手法としてマイクロシミュレーションによる人口予測は、個人属性をもとに人口動態をモデル化するため、小地域での人口予測に優れている。しかし人口予測に必要な世帯属性情報は一般には公表されていないため、推計処理を行う必要がある。鈴木ら¹⁾はアンケート調査を用いて世帯属性情報の推計を行なったが、対象都市ごとにアンケート調査が必要なため、汎用性に欠ける。そこで本研究では、入手可能なオープンデータを用いた初期世帯マイクロデータ推計システムの構築、推定データの妥当性の検証を目的とする。

2. 初期世帯マイクロデータ推定システムの構造

2. 1 システムの概要

本研究により構築した初期世帯マイクロデータ推計システムのフローを図-1に示す。

まず始めに、年齢不詳の個人に対象地域のゾーン、5歳階級、性別人口の割合から年齢データを付与する。次に国勢調査の周辺分布と最初の手順で作成した年齢データを用いて、年齢、性別の属性を持った個人デー

タと世帯人数、世帯タイプ、続柄の属性を持った世帯データを作成し結合する。その後、同ゾーン内の個人同士の入れ替えを行い、入れ替え対象となった世帯内の個人の続柄を年齢の降順に付与し直したあと、誤差関数を求める。誤差関数 $f_L(A)$ は柘井²⁾らの研究で使用されている式(1)を加工して使用する。

$$f(A) = \sum_{i=1}^L \sum_{j=1}^{G_i} |C_{ij}(A) - \text{Round}(M_j(A) * R_{ij}(A))| \quad (1)$$

L は周辺分布の数、 G_i は周辺分布 i に含まれる項目数、 $M_j(A)$ は条件 X を満たす推定中のデータに含まれる人数、 $C_{ij}(A)$ は条件 X, Y を満たす推定中のデータに含まれる人数、 $R_{ij}(A)$ は周辺分布 i の条件 X, Y に含まれる人数の割合を表す。誤差関数の値が小さくなった場合、その入れ替えを採用し次の入れ替えを行っていく。誤差関数が一定回数更新されなかった場合、世帯タイプの入れ替えを複数回行った後、個人への入れ替えに戻る。世帯タイプの入れ替えを行っても誤差関数が更新されない場合、試行回数に応じて変化する確率によって誤差関数が大きくなる入れ替えも採用する。個人、世帯タイプの入れ替え等は全てランダムに行う。

2. 2 対象地域

対象地域として、コンパクトシティ政策を積極的に進めている富山県富山市を選定した。本研究では対象

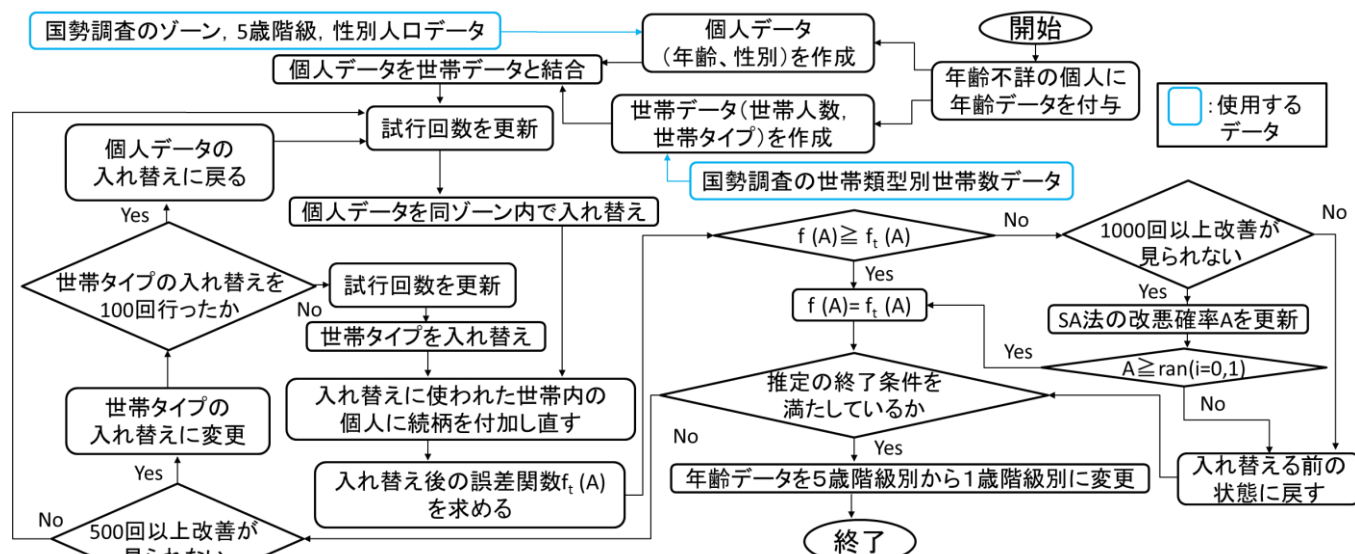


図-1 初期世帯マイクロデータ推定システムのフロー

地域を国勢調査で使用されている 82 ゾーンに分け、初期世帯マイクロデータを推定する。

2. 3 推定に使用する属性

(1) 人口・世帯数

Sugiki et al.³⁾の研究では 7 人世帯までを推定対象とし、7 人以上の世帯を統合していたが、本研究では 8 人世帯を追加し、世帯数と人口の誤差を調整した。

世帯については、国勢調査で使用する家族類型と世帯人数から、世帯人数別の世帯タイプを作成した。構成割合が低いもの、続柄が不明確なものはその他に統合している。使用する世帯タイプを表-1 に示す。

以上から推定値のゾーン、5 歳階級、性別人口及び、ゾーン、世帯人数別世帯数は周辺分布と一致する。

表-1 推定に使用する世帯タイプ

世帯人数	世帯タイプ ()内はコード番号	世帯人数	世帯タイプ ()内はコード番号
1 人	(11)単身・男	4 人	(40)その他
	(12)単身・女	5 人	(51)夫婦+子 3 人
2 人	(21)夫婦のみ		(52)夫婦+両親+子 2 人
	(22)男親+子 1 人		(53)夫婦+子 2 人+親
	(23)女親+子 1 人		(50)その他
3 人	(20)その他	6 人	(61)夫婦+子 4 人
	(31)夫婦+子 1 人		(62)夫婦+両親+子 2 人
	(32)男親+子 2 人		(63)夫婦+子 3 人+親
	(33)女親+子 2 人		(60)その他
	(34)夫婦+親 1 人	7 人	(71)夫婦+子 5 人
4 人	(30)その他		(72)夫婦+両親+子 3 人
	(41)夫婦+子 2 人		(73)夫婦+子 4 人+親
	(42)女親+子 3 人		(70)その他
	(43)夫婦+両親	8 人	(80)その他
	(44)夫婦+子+親	-	(90)施設等世帯

(2) 年齢

年齢階層は 5 歳階級別とし、15 歳未満と 85 歳以上については年齢 1 と年齢 16 にそれぞれ統合した。

2. 4 周辺分布

推定に使用する周辺分布として国勢調査の夫の年齢、妻の年齢別夫婦数のデータと、鈴木ら¹⁾の研究で用いられている出生確率モデルから、親と子供の年齢差を算出したデータの二つを用いる。

3. 推定結果

推定結果の一部を表 2、図 2、図 3 に示す。図 2、表 2 より推定値と周辺分布、出生確率モデルの値との誤差は概ね小さい結果となった。以上から本モデルによって、妻と子供の年齢差、夫と妻の年齢差の二つを考慮した世帯マイクロデータを推定することができたと考えられる。また図 3 より誤差の減少経過は、誤差が小さくなるにつれて改善効率が下がる結果となった。

表-2 妻の年齢に対する夫の年齢の割合 誤差 0.1%以下

		妻の年齢					
		35-39	40-44	45-49	50-54	55-59	60-64
夫の年齢	35-39	52.99%	10.06%	1.11%	0.08%	0.02%	0.01%
	40-44	27.33%	46.37%	6.34%	0.49%	0.10%	0.01%
	45-49	6.88%	32.44%	42.80%	3.83%	0.48%	0.09%
	50-54	1.81%	7.20%	38.18%	38.50%	2.68%	0.29%
	55-59	0.56%	2.01%	9.14%	46.27%	38.43%	3.91%
	60-64	0.28%	0.74%	1.87%	10.13%	52.52%	44.48%
	その他	10.15%	1.18%	0.56%	0.69%	5.79%	51.22%

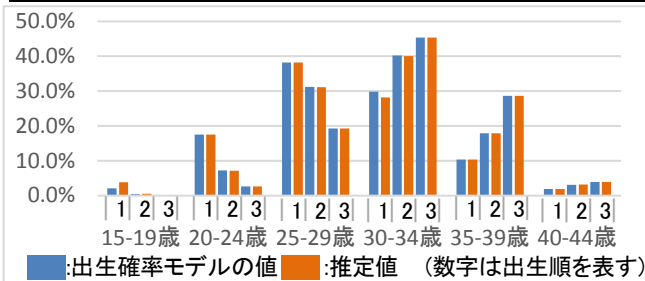


図-2 妻と子供の年齢差、出生順別世帯数の割合

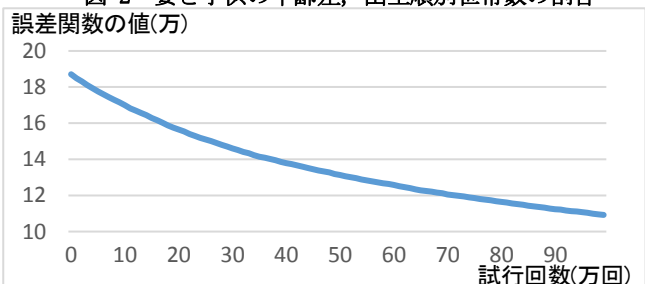


図-3 誤差の減少の途中経過

4. おわりに

本稿では初期の世帯マイクロデータの推定システムの流れとその結果の一部を示した。今後はその他の推定モデルとの比較、及びその他の周辺分布との比較を行い、本モデルの有意性を検証する。

謝辞

本研究の一部は、科学研究費補助金 (17K06597、代表：杉木直) の助成を受けたものである。

参考文献

- 1) 鈴木温・杉木直・宮本和明：空間的マイクロシミュレーションを用いた都市内人口分布の将来予測—人口 40 万人規模の富山市を対象として—，都市計画論文集 Vol.51 No.3，2016 年
- 2) 栢井大貴・村田忠彦：統計データからの市民の属性復元のための進化計算と SA による 2 段階最適化，システム制御情報学会論文誌，Vol.30，No.6，2017 年
- 3) Nao Sugiki, Tomoya Muranaka, Noriko Otani, Kazuki Miyamoto: Agent-based Estimation of Household Microdata with Detailed Attributes for a Real City, Proceedings of the 14th International Conference on Computers in Urban Planning and Urban Management, pp.231-1 - 231-18 2015 年