

特集：第8回人口移動調査の結果から（その1）

移動経歴と初婚発生に関するライフコース分析 — 系列分析（最適マッチング分析・回帰木分析）による類型化 —

鎌田健司・小池司朗・山内昌和*

本稿は「第8回人口移動調査」のデータを用いて、人口移動と家族形成行動の関係性を視覚化・類型化することを目的として、調査時点で40歳以上の男女における15歳から40歳までの移動経歴と初婚発生の関係について系列分析を行った。

その結果、少子化を牽引する世代である1960～70年代生まれは、それ以前の世代に比べて、大都市圏出身者の割合が比較的高く、かつ晩婚化、未婚化が進展していることから、年齢別にみたライフコースの多様化が進んでいることがわかった。回帰木分析による類型化の結果、上位2位の類型は男女ともに1940～50年代以前生まれの「皆婚」世代の移動経歴・初婚発生パターンであった。3位以降の類型では男女で差がみられ、男性は「地元就職・Uターン就職型」で比較的学歴が低く未婚率が高い類型が上位に位置する一方で、女性は比較的学歴が高く未婚率も高い「就職時移動、同・異他県就職型」の移動類型が上位に位置するといった違いがみられた。

このように、ライフコース全体を一つのデータとして分析する系列分析は、これまでの分析では十分に明らかにされてこなかったライフコースの多様性という観点から視覚化、類型化を行うことができる分析手法として有用性が示された。

I. はじめに

政府による地方創生の取り組みが始まり5年が経とうとしている。地方での仕事作りや移住促進、地域振興など様々な取組を行っているものの、東京圏の一極集中の是正や国民の希望出生率の達成などは実現されておらず、人口分野における成果は乏しい状況が続いている。人口移動は就学や就職（転勤）、住宅取得、結婚・出生、家族の随伴移動などライフコース上のイベントと同時または時間的ラグをもって生じ、特に人口移動と結婚や出生といった家族形成行動の関係解明は、東京圏の一極集中や少子化状況の克服のための方策を考える上で重要な研究課題であるといえよう。

人口移動と家族形成行動の関係性を分析する手法としては、例えば初婚行動の時間経過と発生確率を対象とするイベントヒストリー分析による研究が一般的である。しかし、イベントヒストリー分析は1つのイベントに着目して時間経過と生起確率に着目する点で優れた分析手法であるが、競合する複数のイベントを含めたライフコース全体の分析を行う

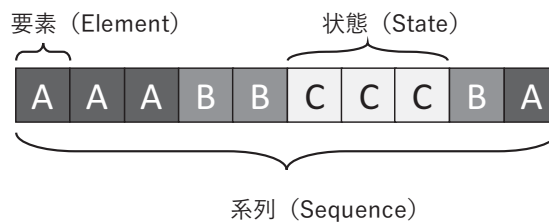
* 早稲田大学教育・総合科学学術院

のには不向きである (Billari 2005). そこで、本稿ではライフコース全体の観察に焦点を当て、人々のライフコースの軌跡を記述・分析する方法である系列分析 (Sequent Analysis) を用いて、人口移動と結婚行動に関するライフコースを視覚化、類型化することを目的とした分析を行う。

系列分析は1960年代から心理学、社会学、人口学、疫学などで用いられてきた記述的・説明的な分析手法である (Gabadinho et al. 2009; 2011; 2013). 個人のライフコース全体を分析対象とする点に特徴があり (Billari 2005; Gabadinho et al. 2013), ライフコース全体を一つの系列とみなした視覚化、類型化が研究の主目的となる。人口学においても Billari の研究をはじめとして多くの蓄積がある (Billari 2001; 2005; Billari and Piccarreta 2005; Aassve et al. 2007; Barban and Billari 2012; Fulda 2016など). また、系列分析は個人のライフコースの「多様性」(diversity) を理解することに着目し、個人の社会経済的属性や移動経歴等の属性によるライフコースの軌跡の違いを多様性の変化として分析することが可能である (Elder et al. 2003).

系列分析では「系列データ」(Sequence data) という特殊な形式のデータを用いる。系列データは「要素」(Element) ・「状態」(State) ・「系列」(Sequence) の3つの情報を持つデータである (図1). 「要素」は時間で区切られた単位を示し、「状態」はその要素の属性を示す。そして状態が記述された要素を時系列の順番で配置したものを「系列」といい、ある時間軸における系列の全体を系列データと呼ぶ。基本的に、時系列上でデータが構成されているため、時系列データ (time series data) やイベントヒストリー分析で用いられる生存時間データ (survival time data) の形式と似た構造ではあるが、系列分析は系列データ全体を観察するため、イベントヒストリー分析における状態変化に関するハザードや観察打ち切りを示すセンサリングの概念はなく、データの並び方それ自体が分析対象であるという点に特徴がある。

図1 系列データの構造と名称



本稿では、国立社会保障・人口問題研究所が2016年に実施した「第8回人口移動調査」のデータを用いて、人口移動と家族形成行動の関係性を系列分析により視覚化・類型化することを目的とする。データ利用上の理由から、人口移動は都道府県間移動のみを対象とし、家族形成行動は初婚行動に限定した分析を行う。研究計画は、第1に人口移動と初婚行動のイベント発生を状態系列とみなした系列データを作成し、人口学的属性や移動経歴によるライフコースの視覚化を行う。第2に、最適マッチング分析によるライフコース間

の距離行列を算出し、機械学習の一手法である回帰木分析によるライフコースの類型化ならびに男女差の分析を行う。

II. データと分析手法

1. データ

使用するデータは、国立社会保障・人口問題研究所が2016年に実施した「第8回人口移動調査」である（国立社会保障・人口問題研究所 2018）。対象世帯数は67,098世帯、有効調査票回収世帯数は48,477世帯（有効回収率は72.2%）、有効回答を得た世帯員は122,640人である。なお、2016年4月に発生した熊本地震の影響により、熊本県および大分県由布市は調査対象から外れた。

本分析では15歳から40歳までを観察期間とし、ライフコースに関する情報が得られる40歳以上の男女を分析対象とした。その結果、59,210ケースが分析対象となる（男性28,191ケース、女性31,029ケース）。

(1) 人口移動（移動経歴）の定義

人口移動に関する系列データを作成するにあたり、問13（生まれた場所）、問14（中学校を卒業したときの居住地）、問15（最後の学校を卒業したときの居住地）、問16（はじめて仕事をもったときの居住地）、問17（はじめての結婚の直前の居住地）、問18（はじめて結婚した直後の居住地）、問19（5年前の居住地）、問20（1年前の居住地）の情報を用いた。これらの変数は、現住地との関係から、同都道府県内であれば市区町村内・外への移動情報が得られるが、出生地との関係から各時点の移動情報を作成するためには都道府県間移動を対象とせざるを得ない。

そのため、人口移動は都道府県間移動に限定し、各時点の居住地を「大都市圏」と「非大都市圏」に分け、出生地との関係から（1）出生地と同じ非大都市圏、（2）出生地と別の非大都市圏、（3）出生地と同じ大都市圏、（4）出生地と別の大都市圏、の4カテゴリに分けて分析を行う。なお、大都市圏—非大都市圏の定義は都道府県別に行い、大都市圏は東京都、埼玉県、千葉県、神奈川県、愛知県、岐阜県、三重県、大阪府、京都府、兵庫県が該当し、それ以外の道県は非大都市圏とした。

(2) 家族形成行動（初婚発生）の定義

家族形成行動については、データの利用上、初婚の発生のみを対象とした。本調査では、結婚時期に関する情報は問17において「初婚時の年齢」が得られる。ただし、初婚が離別・死別によって解消されている場合に、離別年齢、死別年齢が得られないことから、今回は初婚が発生した場合、現在の配偶関係が離別・死別であっても、初婚状態が継続すると仮定した系列データを作成した。また、現在の配偶関係が有配偶、離別、死別で初婚年齢が不詳の4,123ケースについては、性別、調査時年齢、最終学歴を独立変数とした回帰モデ

ルを用いた多重代入法によるデータの代入を行った¹⁾。

(3) 観察する属性情報

本分析で用いる属性情報は、人口学的属性である出生年、出生順位、最終学歴、出生地域、経済学的属性である初職、移動経歴を示す「学卒・初職地類型」、「出生地・現住地類型」である。

移動経歴である「学卒・初職地類型」は林（2015）を参考に作成した。この類型は最終学卒時・初職時の都道府県と出生都道府県との関係から5類型及び類型不詳の6類型が得られる。

- 1) 「類型11-1(地元就職)：学卒・初職時が出生地と同じ(11)であり、学卒・初職時と同じ(1)」
- 2) 「類型10-0(就職時移動)：学卒時は出生地と同じ(1)だが、初職時は異なり(0)、学卒時と初職時も異なる(0)」
- 3) 「類型01-0(Uターン就職)：学卒時は出生地と異なる(0)が、初職時は同じ(1)であり、学卒時と初職時は異なる(0)」
- 4) 「類型00-1(同他県就職)：学卒時・初職時は出生地と異なり(00)、学卒時と初職時は同じ(1)」
- 5) 「類型00-0(異他県就職)：学卒時・初職時は出生地と異なり(00)、学卒時と初職時も異なる(0)」
- 6) 「類型不詳」

次に、「出生地・現住地類型」は小池（2006; 2009; 2014）を参考に作成した。この類型は出生時と現在における居住地を非大都市圏(0)-大都市圏(1)別に分けて4類型を構成する。大都市圏の定義は移動経歴の定義と同様である。

- 1) 「類型b0c0(RR)：出生地(b)と現住地(c)が非大都市圏(0)」
- 2) 「類型b0c1(RU)：出生地(b)が非大都市圏(0)、現住地(c)が大都市圏(1)」
- 3) 「類型b1c0(UR)：出生地(b)が大都市圏(1)、現住地(c)が非大都市圏(0)」
- 4) 「類型b1c1(UU)：出生地(b)と現住地(c)が大都市圏(1)」

2. 分析手法

(1) ライフコースの多様性指標 (Entropy Index)

系列分析において、ライフコースの性質を示す中心的な役割を持つ指標が多様性指標

1) 初婚年齢の不詳割合は、有配偶（同居）6.5%、有配偶（別居）9.4%、離別11.2%、死別12.4%であった。推定モデルは、 $\text{初婚年齢} = 26.58 + \text{性別} * 2.6 - \text{調査時年齢} * 0.04 + \text{最終学歴（高校卒）} * 0.55 + \text{最終学歴（短大・高専卒）} * 1.48 + \text{最終学歴（大学・院卒）} * 2.37$ である。その結果、データ代入前の「現在の配偶関係」別、平均初婚年齢は、有配偶者（同居）39,983ケース26.45歳、有配偶者（別居）1,104ケース26.52歳、離別2,941ケース25.46歳、死別5,964ケース23.98歳、代入後は、有配偶者（同居）42,782ケース26.41歳、有配偶者（別居）1,218ケース26.45歳、離別3,310ケース25.48歳、死別6,805ケース23.91歳であり、代入前後で分布の差は生じていない。多重代入法にはStata 14.2のmi関数を用いた。

(Entropy Index) である。この指標は、情報理論で用いられるシャノンの平均情報量 (Shannon 1948) を援用している。シャノンの平均情報量は、一般的に「エントロピー」、 「平均情報量」、 「シャノン情報量」と称されるが、ライフコース分析においては、「状態の多様性」 (“diversity of states”, Gabadinho et al. 2011) として解釈することができる。本稿においてはシャノンの平均情報量を「多様性指標」として統一して表記する。

多様性指標は以下のように定式化される。

$$h(p_1, \dots, p_a) = - \sum_{i=1}^a p_i \log(p_i)$$

p_i : 状態 (state) i の割合, a : 状態の数

多様性指標は、すべての状態の確率の総和が1になる完全事象系において、各状態の構成割合の変化によって値が変化するという性質を持つ。あるイベントに複数の状態 (states) があるとき、一つの状態に収斂すると情報量は0となり、すべての状態が同じ構成割合になると最大値 $\log_2 a$ となる。すなわち、各状態の構成割合が均一化すると、多様性指標の数値が高くなる (ライフコースの多様性が増えると解釈)。従ってこの数値はイベント発生確率を示す指標ではないことに留意が必要である。

多様性指標には2つの側面から観察が可能である。第1に年齢別に複数ある状態間の構成割合の変化を多様性として観察する方法である (年齢別にみた多様性)。これは、ある年齢における個人間 (Between sequence entropy) のライフコースの多様性を示す。本稿では15~40歳の各年齢において、属性別の多様性指標を算出する。第2の方法は、個人のライフコースの多様性を観察する方法である (個人の多様性)。これは個人単位で観察され、個人内 (Within sequence entropy) のイベントの状態変化をライフコースの多様性として解釈するものであり、本稿では15~40歳の期間における各状態の構成割合から、単一の多様性指標が算出される。すなわち、15~40歳の期間全体において、複数あるイベントの状態の構成割合が均一に近づくと多様性指標は高くなり、同一の状態の構成割合が高くなると多様性が低くなる。このように個人単位で観察される個人の多様性指標を回帰木分析に用いることで、ライフコースの類型化を行う。なお、個人の多様性指標は分析モデル間の解釈を容易にするために正規化を行った値を用いる。

各モデルの個人の多様性指標の記述統計を表1に示した。移動経歴の多様性指標の平均値は男性0.128、女性0.108、最大値は男性0.791、女性0.788、初婚発生の平均値は男性0.817、女性0.826、最大値は男女ともに1である。移動経歴・初婚発生の平均値は男性0.311、女性0.300、最大値は男性0.714、女性0.665である²⁾。

2) 本分析は Gabadinho et al. (2011) が開発した R のパッケージ (関数群) である TraMineR を用いて主な分析を行った。ソフトウェアは R 3.5.3, TraMineR 2.0-11.1を使用した。多様性指標図には TraMineR の seqHtplot 関数を用い、個人の多様性指標の算出には seqient 関数を用いた。また、状態分布図の出力は seqdplot 関数を用いた。

表 1 個人の多様性指標（移動経歴，初婚発生，移動経歴・初婚発生）の記述統計

	男性 (N=28,181)				女性 (N=31,029)			
	平均値	標準偏差	最小値	最大値	平均値	標準偏差	最小値	最大値
多様性指標								
移動経歴	0.128	0.203	0.000	0.791	0.108	0.190	0.000	0.788
初婚発生	0.817	0.323	0.000	1.000	0.826	0.252	0.000	1.000
移動経歴・初婚発生	0.311	0.125	0.000	0.714	0.300	0.099	0.000	0.665

(2) 状態分布図 (State Distribution Plot) と年齢別にみた多様性指標の推移

系列分析における最も基本的な分析は、時点別・年齢別にライフコースの多様性を視覚化することである。そのための方法として、状態分布図ならびに年齢別にみた多様性指標を示したい。状態分布図は、時点別・年齢別の各状態の構成割合の変化を視覚化する。状態分布図における各状態の構成割合の変化は多様性指標の変化という形で対応するため、年齢別にみたライフコースの多様性を評価するためには両方の分布を照らし合わせて解釈すると理解がしやすい。本稿では紙幅の関係から出生年、最終学歴、学卒・初職地類型の3変数を取り上げる。

(3) 最適マッチング分析 (Optimal Matching Analysis) による距離行列の算出

最適マッチングとは、ある2つのライフコース系列があるときに、両方の系列が同一の系列になる場合の操作量を計算し、その操作量を2つのライフコースの間の距離として定量化する手法である。最適マッチングの考え方は1950年代から情報科学において発展し、1970年代以降は分子生物学におけるDNA配列の分析等に用いられてきた分析手法であり、1980年代にAbbottらによって社会科学に導入された (Abbott and Forest 1986, Abbott 1995, Abbott and Tsay 2000)。日本では職歴パターンの分析 (渡邊 2004)、女性の初期キャリアの類型化 (香川 2010)、女性のライフコースの多様化 (福田 2006) 等の分析事例がある。

最適マッチング分析の目的は、個人のライフコース間の非類似度 (dissimilarity measures) を距離行列として定量化し、均質な系列群を発見することにある (Studer et al. 2011)。この距離行列を用いることによってクラスター分析や回帰木分析による類型化が可能になる。クラスター分析では非類似度を示す距離行列を直接的に用いて対象をクラスターへの分類に用いられるが、回帰木分析においては、距離行列は回帰モデルの当てはまりを示す寄与率 (R^2 値) やモデル検定量であるF統計量の算出に用いられる。従って、系列データを用いた回帰木分析では、最適マッチング分析による距離行列を用いた疑似的な寄与率及びF統計量によって分類・類型化がなされる。

距離行列の算出には「挿入」(insertion)、「削除」(deletion)、「置換」(substitution)の3種類の操作を用いる。系列間の距離は上記の操作量を算出し、各操作に重み付けを行い再計算し、最終的に最も小さい操作量となる操作の組合せを系列間の距離とする。この際、2つの系列の時点数は異なっても問題がない。

一般に距離行列の算出過程において、「置換」は「挿入」や「削除」に比べてコストが小さくなる傾向にある。そのため、「置換」のコストを「挿入」や「削除」よりも高くしたり、状態間の推移確率を用いたりするなど様々な考え方がある。

本稿においては、「挿入」・「削除」コストは1、「置換」コストは状態間の推移率 (transition rate) を割り当てて距離行列の算出を行った。状態間の推移率は $2-p(s_i|s_j)-p(s_j|s_i)$ によって得られる。ここで、 s_i は状態 i 、 s_j は状態 j である。また、最適マッチング分析ならびに回帰木分析は移動経歴・初婚発生モデルについてのみ行った。男女別に最適マッチング分析によって距離行列を算出した結果、男性は7.9億パターン、女性は9.6億パターンの距離行列が得られた³⁾。

(4) 回帰木分析 (Regression Tree)

回帰木は決定木分析 (Decision Tree) の一つであり、定量的な独立変数を用いた線形回帰分析結果をもとに分類を行う分析手法である。回帰木分析はデータマイニング分野の機械学習の一手法として発展してきた探索的かつ記述的なアプローチである (Bresiman et al. 1984)。回帰木分析の伝統的なアルゴリズムである CART (Classification and Regression Tree) は2分木による分類を行う点に特徴があり、分析結果を視覚的かつ理解しやすい形で表現すること、多変量分析が可能になる点などが利点である。ただし、非線形関係をモデル化するとき有効である一方、離散型変数を用いる場合や線形関係の予測力が低いという点などが欠点として挙げられる。

本稿では多変量変数を用いた2分木による回帰木分析を行うことで、男女別の移動経歴・初婚発生モデルの類型化を行いたい。従属変数は移動経歴・初婚発生モデルにおける正規化された個人の多様性指標を従属変数とし、独立変数は出生年、出生順位、最終学歴、初職、学卒・初職地類型、出生地・現住地類型、出生地域とした。記述統計を表2に示した。

系列分析における回帰分析は Studer et al. (2011) のアルゴリズムを用いる⁴⁾。Studer et al. (2011) の回帰木モデルは、基本的に CART のアルゴリズムと同様に2分木法を採用している。回帰木分析の手順は以下の通りである。(1)すべてのケースについて、投入された独立変数を反復的に投入し、最適マッチング分析によって得られた距離行列を用いた疑似寄与率 (R^2 値) が最も高い独立変数を選択し2分類する (親ノードが2つの子ノードを生成)。(2)分類された子ノードにおいて同様の分類を行う。子ノードでは、できるだけ親ノードにおいて選択された独立変数とは異なる変数が選択されるように制約が課される。(3)以下、同様に分類を行い、重回帰分析のモデル検定である疑似 F 統計量が統計学的に有意でなくなった時点で分類が終了し、最終的なライフコースの類型化がなされる。以上のように推定された男女別の移動経歴・初婚発生モデルの回帰木分析結果について、最終的に分類されたノードをケース順に並べた頻出パターンについて男女での比較を行う。

3) 距離行列の算出には TraMineR パッケージの seqdist 関数を用いた。

4) 回帰木分析は TraMineR パッケージの seqtree 関数を用いた。また、回帰木分析の結果を視覚化するソフトウェアである GraphViz 2.38を R から出力するパッケージである Rgrphviz 2.24.0を用いた。

表2 属性情報の記述統計

属性情報	男性 (N=28,181)					女性 (N=31,029)				
	標本数	割合(%)	個人の多様性指標 (平均値)			標本数	割合(%)	個人の多様性指標 (平均値)		
			移動経歴	初婚発生	移動経歴・初婚発生			移動経歴	初婚発生	移動経歴・初婚発生
出生年										
1900-39年	3,868	13.7	0.090	0.934	0.331	5,330	17.2	0.070	0.835	0.289
1940-49年	6,341	22.5	0.139	0.901	0.335	6,867	22.1	0.118	0.851	0.305
1950-59年	6,605	23.4	0.138	0.821	0.317	6,890	22.2	0.121	0.847	0.309
1960-69年	6,566	23.3	0.127	0.745	0.291	6,946	22.4	0.108	0.809	0.296
1970-79年	4,799	17.0	0.131	0.685	0.281	4,996	16.1	0.119	0.770	0.297
出生順位										
第1子	11,213	39.8	0.119	0.803	0.306	12,205	39.3	0.098	0.820	0.296
第2子	11,343	40.2	0.132	0.818	0.313	12,409	40.0	0.112	0.828	0.303
第3子	4,149	14.7	0.140	0.832	0.317	4,684	15.1	0.118	0.832	0.302
第4子以上	1,476	5.2	0.137	0.873	0.324	1,731	5.6	0.113	0.840	0.301
最終学歴										
中学校卒	4,316	15.3	0.075	0.831	0.285	5,218	16.8	0.073	0.808	0.278
高等学校卒	11,040	39.2	0.106	0.806	0.299	13,744	44.3	0.096	0.831	0.297
短大・高専卒	2,592	9.2	0.125	0.791	0.307	3,601	11.6	0.129	0.833	0.311
大学・大学院卒	10,233	36.3	0.196	0.837	0.349	8,466	27.3	0.179	0.820	0.328
初職										
正規職員/役員	23,394	83.0	0.140	0.826	0.318	23,331	75.2	0.118	0.836	0.306
パートタイム・アルバイト	1,279	4.5	0.110	0.685	0.266	2,445	7.9	0.091	0.784	0.282
派遣・嘱託・契約社員	530	1.9	0.117	0.728	0.276	859	2.8	0.116	0.786	0.289
自営・家族従業者・内職	2,535	9.0	0.040	0.851	0.293	2,367	7.6	0.054	0.819	0.282
無職	443	1.6	0.092	0.585	0.228	2,027	6.5	0.085	0.795	0.280
学卒・初職地類型										
類型11-1(地元就職)	16,781	59.5	0.035	0.813	0.275	22,155	71.4	0.057	0.826	0.283
類型10-0(就職時移動)	4,085	14.5	0.361	0.820	0.393	2,921	9.4	0.339	0.836	0.374
類型01-0(Uターン就職)	1,778	6.3	0.098	0.851	0.315	1,263	4.1	0.147	0.838	0.317
類型00-1(同他県就職)	3,323	11.8	0.247	0.798	0.355	3,609	11.6	0.211	0.813	0.336
類型00-0(異他県就職)	1,723	6.1	0.378	0.822	0.420	623	2.0	0.309	0.809	0.380
類型不詳	491	1.7	0.077	0.831	0.291	458	1.5	0.088	0.844	0.295
出生地・現住地類型										
類型b0c0(RR)	12,790	45.4	0.115	0.820	0.309	14,365	46.3	0.091	0.825	0.297
類型b0c1(RU)	3,803	13.5	0.274	0.842	0.367	4,012	12.9	0.261	0.847	0.341
類型b1c0(UR)	952	3.4	0.226	0.798	0.330	1,056	3.4	0.209	0.840	0.321
類型b1c1(UU)	10,636	37.7	0.108	0.802	0.298	11,596	37.4	0.102	0.818	0.293
出生地域										
北海道・東北地方	4,244	15.1	0.121	0.808	0.307	4,601	14.8	0.090	0.817	0.295
関東地方	7,198	25.5	0.126	0.786	0.298	8,259	26.6	0.121	0.823	0.300
中部地方	5,738	20.4	0.104	0.821	0.307	6,175	19.9	0.087	0.842	0.301
近畿地方	4,239	15.0	0.136	0.828	0.316	4,511	14.5	0.122	0.826	0.300
中国・四国地方	3,138	11.1	0.132	0.828	0.318	3,460	11.2	0.108	0.826	0.301
九州地方・沖縄県	3,625	12.9	0.158	0.826	0.323	4,023	13.0	0.131	0.819	0.302

Ⅲ. 結果

1. 状態分布図 (State Distribution Plot) ・多様性指標

状態分布図は各年齢における状態の構成割合の変化を示し、多様性指標はその構成割合の変化を「多様性」として表している。図2～4には、出生年、最終学歴、学卒・初職地類型の3属性について (a)状態分布図および (b)多様性指標の年齢別の推移を示している。以下では (b)年齢別にみた多様性指標の推移を中心に説明する。

(1) 移動経歴

移動経歴は各時点の居住地を「大都市圏」と「非大都市圏」に分け、出生地との関係から (1)出生地と同じ非大都市圏、(2)出生地と別の非大都市圏、(3)出生地と同じ大都市圏、(4)出生地と別の大都市圏の4カテゴリに分けられた属性情報である。

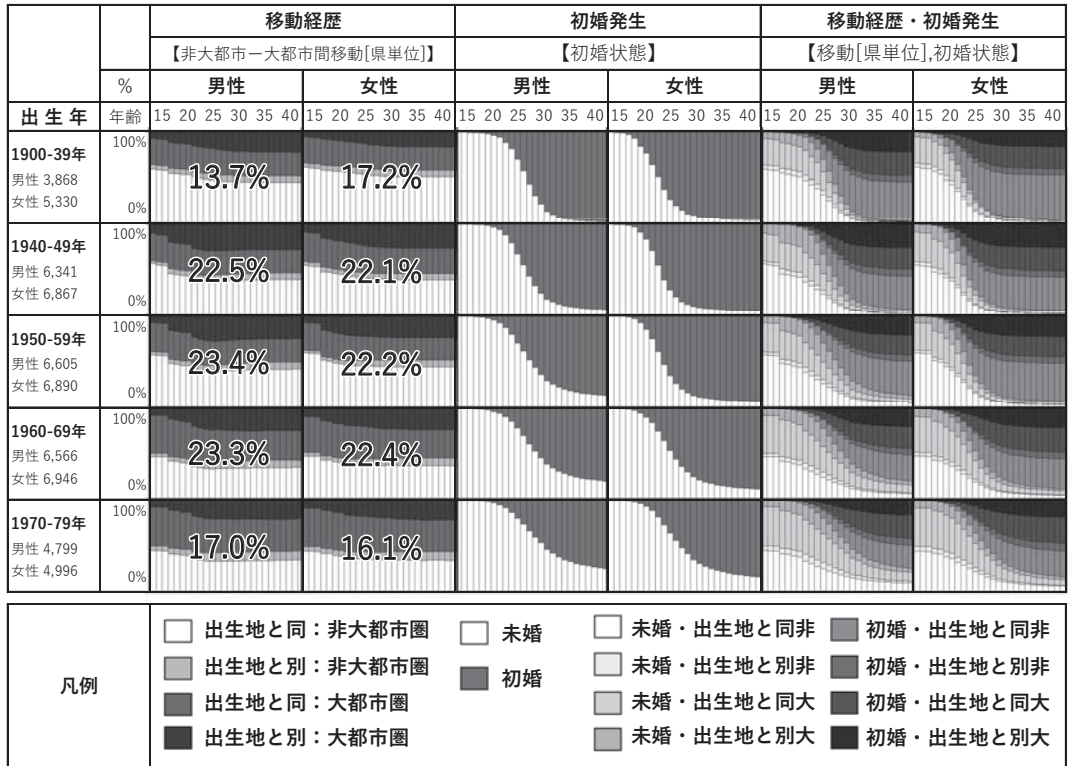
出生年別に年齢別の多様性指標の推移をみると (図2)、1940年代～1960年代生まれの世代では、18歳時点、20歳前半に多様化指標の上昇がみられる。これは最終学歴卒業時点もしくは初職時点での移動経歴の構成割合の変化が生じており、大都市圏への移動割合が増加していることを示す。一方で1970年代生まれは、非大都市圏出身者の減少・大都市圏出身者の増加、出生地とは別の地域での居住の増加を受けて多様性は高い水準にある。

最終学歴別に年齢別の多様性指標の推移をみると (図3)、非大都市圏から大都市圏への移動と強く結びついているという傾向が顕著にわかる。移動時期については、人口移動調査では卒業時点の居住地域から移動の系列データを作成しているため、卒業時点で移動が生じていることに留意が必要である。「中学校卒」(男性15.3%、女性16.8%)は15歳から移動が開始され、「高等学校卒」(男性39.2%、女性44.3%)は18歳から、「短大・高専卒」(男性9.2%、女性11.6%)は18～20歳前半に段階的に移動が生じ、「大学・院卒」(男性36.3%、女性27.3%)では21歳時点で大きな変化が生じている。

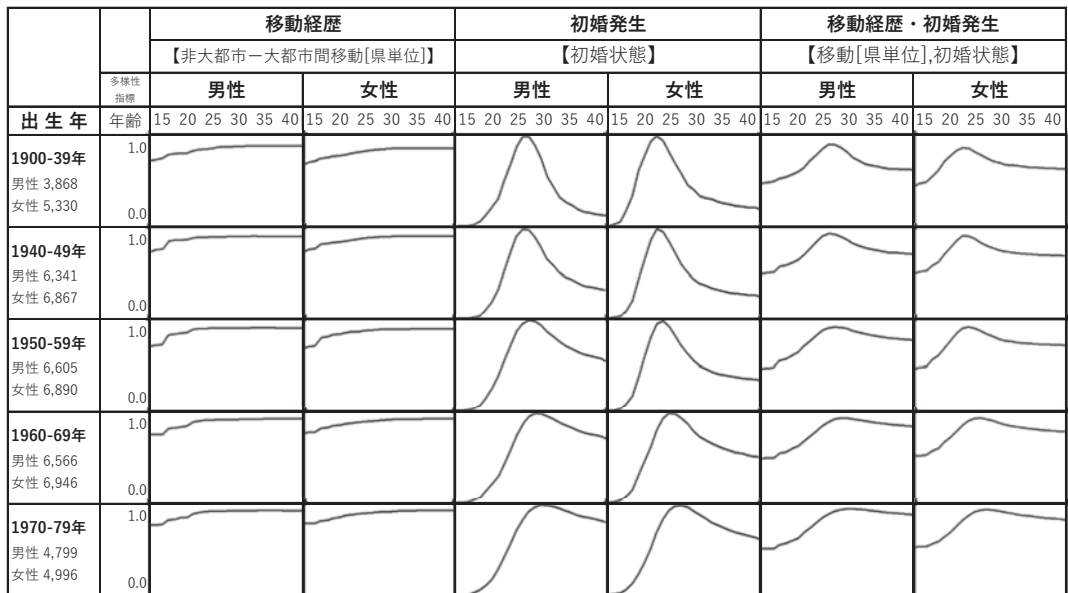
学卒・初職地類型別に年齢別の多様性指標の推移をみると (図4)、「地元就職」(男性59.5%、女性71.4%)は大半が出生地と同じ非大都市圏・大都市圏での継続的な居住が観察される。「就職時移動」(男性14.5%、女性9.4%)は18歳での変化が大きく、大半は出生地とは別の大都市圏への移動であるが、出生地と別の非大都市圏への移動も一定数含まれる。年齢を経るとともに出生地と同じ非大都市圏居住の構成割合が増えることから多様性指標の上昇がみられる。「Uターン就職」(男性6.3%、女性4.1%)は非大都市圏出身者に多く、大都市圏出身者でも一定数が存在し、Uターン後に出生地とは別の大都市圏へ移動するケースも多く、男性よりも女性の方が多い。「同他県就職」(男性11.8%、女性11.6%)は20代前半での移動が多く、その移動先は出生地と別の大都市圏および非大都市圏が多いことから、短大・高専卒や大学・院卒で大都市圏へ移動、居住する類型である。「異他県就職」(男性6.1%、女性2.0%)は20代前半での出生地と別の大都市圏・非大都市圏への移動が顕著であり、短大・高専卒や大学・院卒で大都市圏へ移動し、その後出生地

図2 出生年別、移動経歴・初婚発生についての状態分布図および多様性指標

(a) 状態分布図



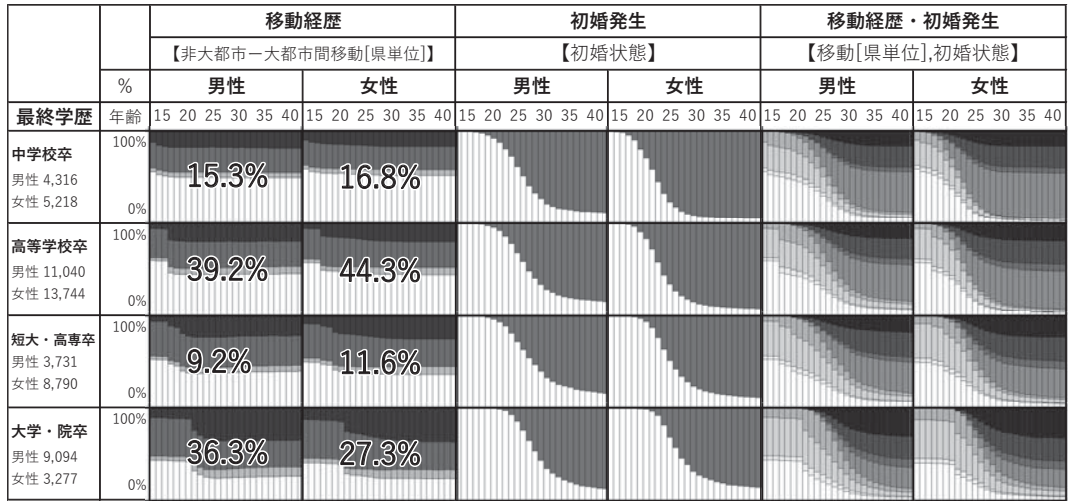
(b) 多様性指標



(注) 状態分布図内の数値は、各カテゴリの構成割合を示す。

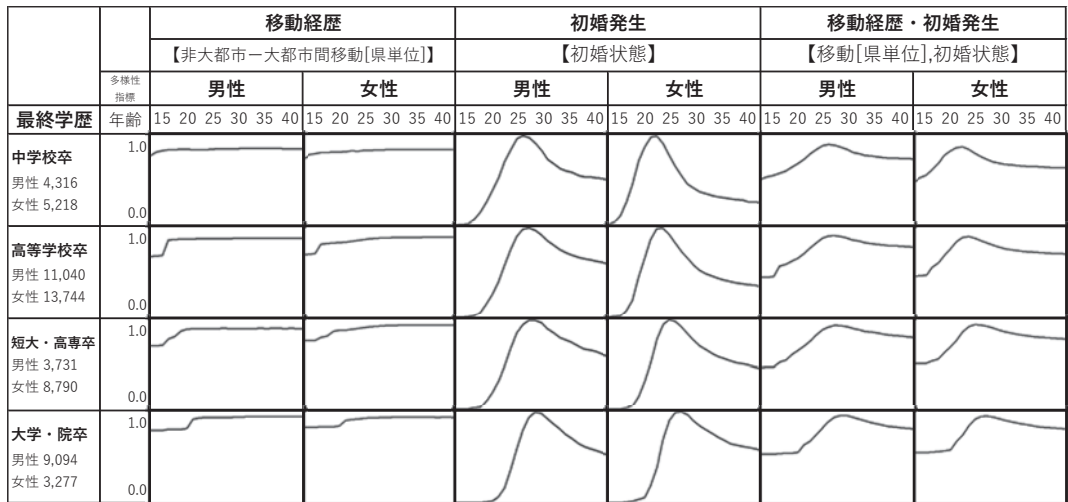
図3 学歴別，移動経歴・初婚発生についての状態分布図および多様性指標

(a) 状態分布図



凡例は省略（図2を参照）

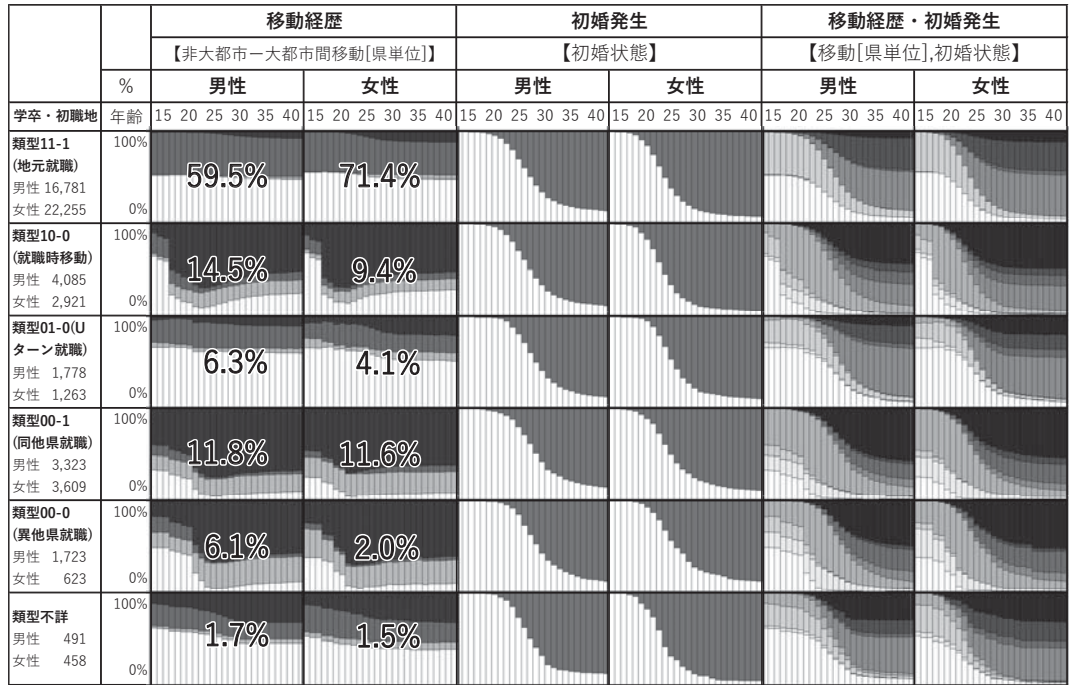
(b) 多様性指標



(注) 状態分布図内の数値は，各カテゴリの構成割合を示す。

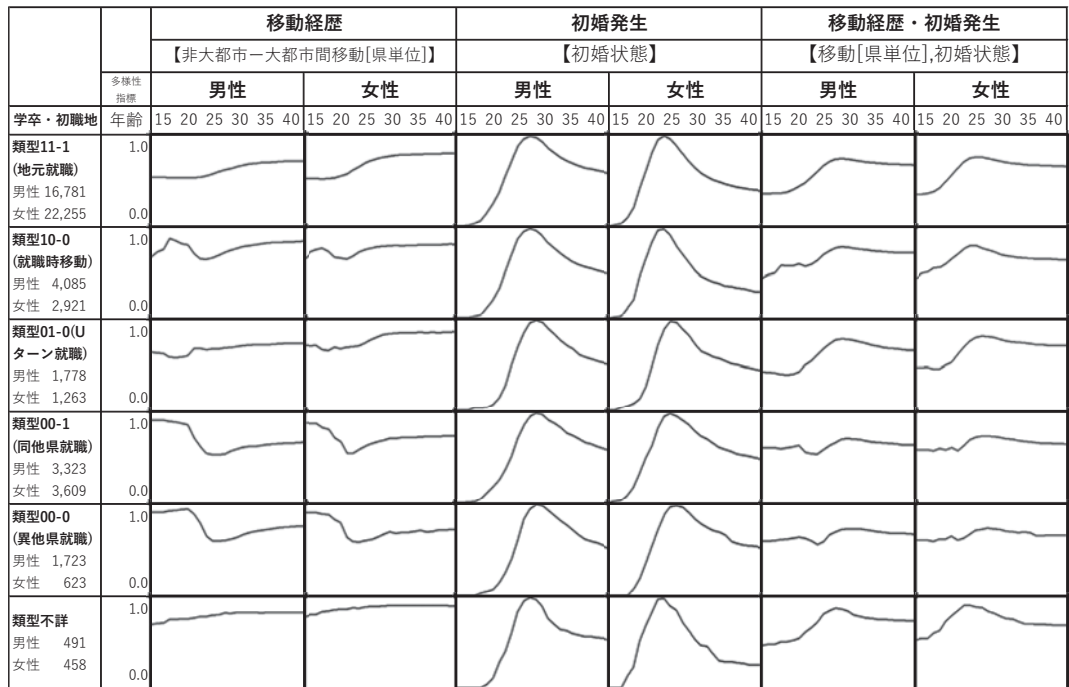
図4 学卒・初職地類型別，移動経歴・初婚発生についての状態分布図および多様性指標

(a) 状態分布図



凡例は省略（図2を参照）

(b) 多様性指標



(注) 状態分布図内の数値は，各カテゴリの構成割合を示す。

とは別の大都市圏・非大都市圏で初職を得ている類型である。「同他県就職」や「異他県就職」は学卒時に出生地と異なる地域へ移動が生じるが、その後は居住地域の構成割合は安定的であるため多様性指標は低い水準となる。

(2) 初婚発生

初婚発生における年齢別にみた多様性指標は、すべての人が15歳時点では未婚状態であることから0から始まり、初婚経験者が増加するに従い多様性指標の数値は上昇し、未婚割合＝初婚経験割合となる時点で最大値である1になる。その後、初婚経験割合が50%を上回って上昇することによって多様性指標は低下するという傾向を示す。年齢別にみた多様性指標において、40歳時点での未婚者割合の増加、すなわち未婚化の進展は未婚状態と初婚経験の構成割合を均一化させる方向に作用するため多様性指標の数値が上昇するという性質をもつ。

出生年別に年齢別の多様性指標の推移をみると(図2)、出生年が近年になるに従って晩婚・未婚化していることがわかる。晩婚化はピーク時の年齢の高年齢へのシフトで把握することができ、未婚化はピーク以後の多様性指標の低下幅が縮小することで示される。男女差では女性よりも男性の方が晩婚傾向、40歳時点での未婚者割合が高い。

最終学歴別にみると(図3)、中学校卒から大学・院卒になるにつれて男女ともに晩婚傾向が進み、多様性指標のピークが高年齢にシフトする。ただし、男性は大学・院卒の40歳時点での未婚者割合は他の学歴に比べて低い傾向にあり、対照的に女性は大学・院卒の方が他の学歴に比べて未婚傾向にある。

学卒・初職地類型別に年齢別の多様性指標の推移をみると(図4)、「地元就職」の男性は他の類型に比べて早婚傾向にあるが、40歳時点の未婚者割合は相対的に高く、女性は早婚傾向かつ40歳時点での未婚者割合は低い。18歳での移動が多い「就職時移動」の男性は比較的早婚傾向かつ40歳時点での未婚者割合も低い。女性も同様に早婚傾向であり、40歳時点での未婚者割合も低い。「Uターン就職」は男女ともに他の類型と比べて平均的な初婚経験傾向および40歳時点での未婚者割合である。「同他県就職」・「異他県就職」は大学・院卒を含む移動類型であることもあり、晩婚・未婚化傾向が顕著である。

(3) 移動経歴・初婚発生

移動経歴・初婚発生はこれまでみてきた移動経歴と初婚発生の特性をクロスさせた8カテゴリである(図2～4)。基本的な傾向は移動経歴と初婚発生の特徴を兼ねたものであるため詳述はせず、最頻出系列(modal state sequence)に着目してみたい。

出生年別に年齢別の多様性指標の推移をみると(図2)、1940年代から1950年代生まれの世代は18歳時点での移動が顕著に観察され、出生年が近年になるほど男女ともに晩婚・未婚化が生じている。出生年別にみた最頻出系列は、1950年代生まれまでは、基本的に男女ともに出生地と同じ非大都市圏において20代で初婚を経験する者である。1960年代生まれになると、男女ともに10代から20代の未婚期においては出生地と同じ大都市圏居住者が

最も多くなる。1970年代生まれは、10代後半から出生地が大都市圏である者が最も多く、初婚経験後においては、出生地と別の大都市圏居住者の割合が最も高い。

最終学歴別に年齢別の多様性指標の推移をみると（図3）、各学歴における卒業時点で移動が生じ、その後は高学歴ほど晩婚化傾向を示す。40歳時点での未婚者割合については、男性は大学・院卒で比較的低く、女性は比較的高い点が対照的である。最終学歴別にみた最頻出系列は、男女ともに中学校卒、高等学校卒では出生地と同じ非大都市圏において初婚経験者である。大学・院卒は卒業時期である20代前半から出生地と別の大都市圏居住者が最多となり、初婚経験後も出生地と別の大都市圏居住者が最も多い。

学卒・初職時類型別に年齢別の多様性指標の推移をみると（図4）、学卒時の移動傾向と最終学歴に対応した初婚経験傾向を示している。「地元就職」「Uターン就職」は男女ともに出生地と同じ非大都市圏で初婚経験をする者が最も多い。「就職時移動」は男女ともに18歳までは出生地と同じ非大都市圏居住者が最も多く、19歳から20代後半までは出生地と別の大都市圏で居住する初婚経験者が最も多い。「同他県就職」は15歳時点で出生地と別の大都市圏に居住する初婚経験者が最も多く、「異他県就職」は10代では出生地と同じ非大都市圏に居住しているが、その後大都市圏に居住する初婚経験者が多い。

2. 回帰分析結果

(1) 男性の移動経歴・初婚発生モデル

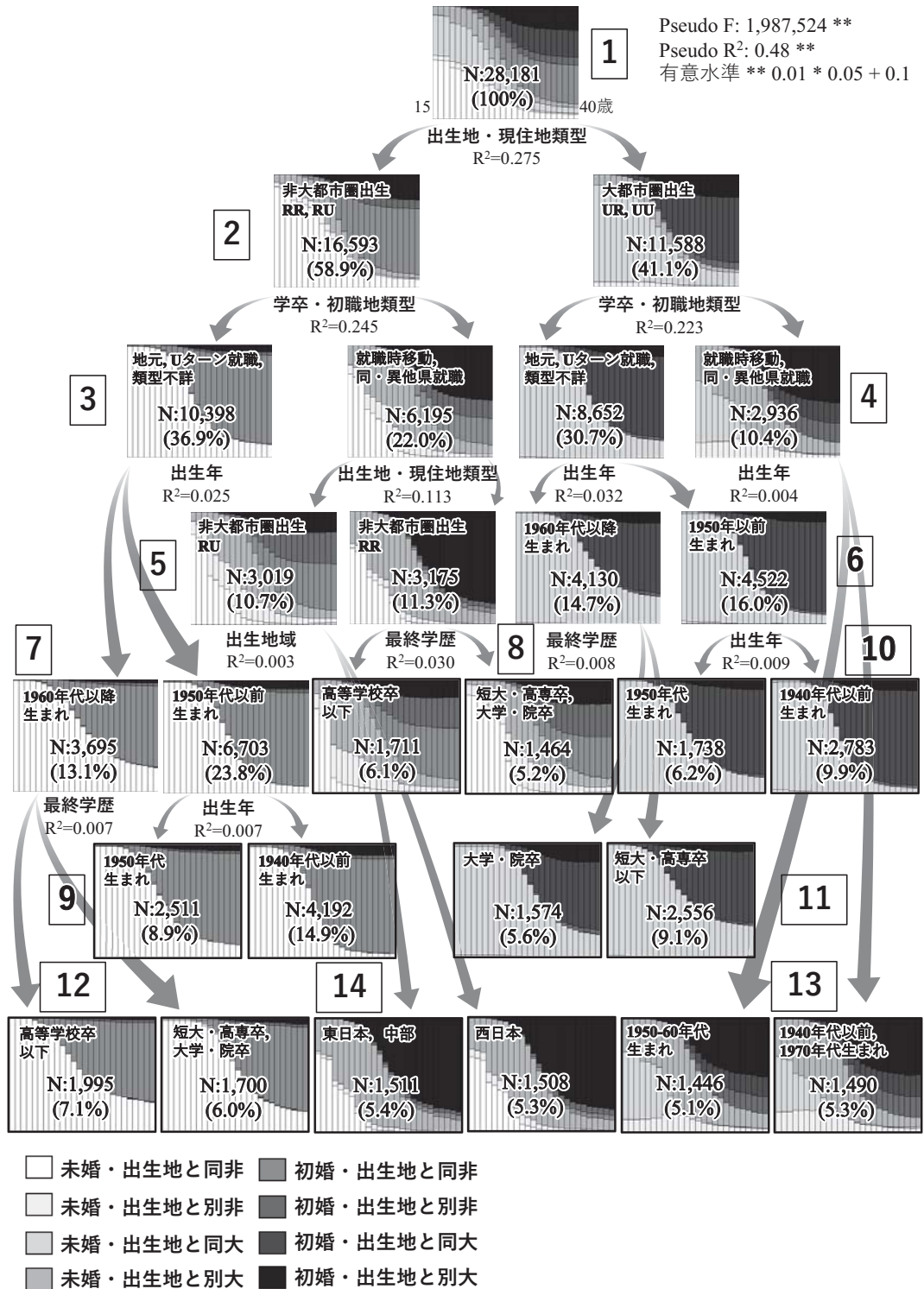
男性の移動経歴・初婚発生モデルについて回帰分析を行った結果、14階層の分類がなされた（図5）。男性モデルの疑似寄与率（ R^2 値）は0.48であった。図中の数字は分類の順番を示すものである。系列には分類されたカテゴリの略称とケース数、総数に占める割合を示している。

全体的な分類の傾向をみると、第1に出生地（「非大都市圏」（RR, RU, 58.9%）、「大都市圏」（UR, UU, 41.1%））によって移動経歴・初婚発生パターンが大きく分かれる（第2階層）。次に学卒・就職時期における移動類型で分類され、非大都市圏出生群と大都市圏出生群ともに「地元就職、Uターン就職、類型不詳」群（非大都市圏36.9%、大都市圏30.7%）と「就職時移動、同・異他県就職」群（非大都市圏22.0%、大都市圏10.4%）に分けられる（第3・4階層）。前者は出生地と同じ地域で就職する群であり、後者は出生地と異なる地域（主に大都市圏）で就職する群である。その後のライフコースの違いは主に出生年と最終学歴の違いによる分類がなされる。

それでは第5階層以下の分類についてみていきたい。非大都市圏出生の「地元就職、Uターン就職、類型不詳」群は、出生年によって「1950年代以前生まれ」（23.8%）と「1960年代以後生まれ」（13.1%）に分類され（第7階層）、1960年代以降生まれ世代において晩婚・未婚化傾向が顕著になる。さらに1960年代以降生まれは、最終学歴によって「高等学校卒以下」（7.1%）と「短大・高専、大学・院卒」（6.0%）に分類される（第12階層）。前者は後者よりも早婚傾向であるが40歳時点での未婚者割合が高い傾向を示す。

一方で、非大都市圏出生の「就職時移動、同・異他県就職」群は、「非大都市圏出生で

図5 回帰分析による、男性の移動経歴・初婚発生モデルの推定結果



現在も非大都市圏に居住する」群 (RR, 11.3%) と「非大都市圏出生で現在は大都市圏に居住する」群 (RU, 10.7%) に分けられる (第5階層)。「RR」群は、さらに最終学歴によって「高等学校卒以下」(6.1%) と「短大・高専, 大学・院卒」(5.2%) に分類される (第8階層)。前者は18歳において移動が生じる群, 後者は20代前半に移動が生じる群である。「RU」群は、出生地域によって「東日本, 中部」地方 (5.4%) と「西日本」(5.3%) に分類される (第14階層)。「東日本, 中部」地方は18歳時点の大都市圏への移動が西日本よりも大きい変化が生じている一方で、「西日本」は出生地と異なる非大都市圏で居住する割合が東日本, 中部地方よりも高い。

大都市圏出生群においても、出生年と最終学歴によって分類される。「地元, Uターン就職, 類型不詳」群は、出生年によって「1950年以前生まれ」(16.0%) と「1960年代以降生まれ」(14.7%) で分類され (第6階層), 後者は晩婚・未婚化の傾向を示す。「1960年代以降生まれ」は、さらに最終学歴によって「大学・院卒」(5.6%) と「短大・高専卒以下」(9.1%) に分類される (第11階層)。前者の40歳時点での未婚者割合は後者よりも低い。

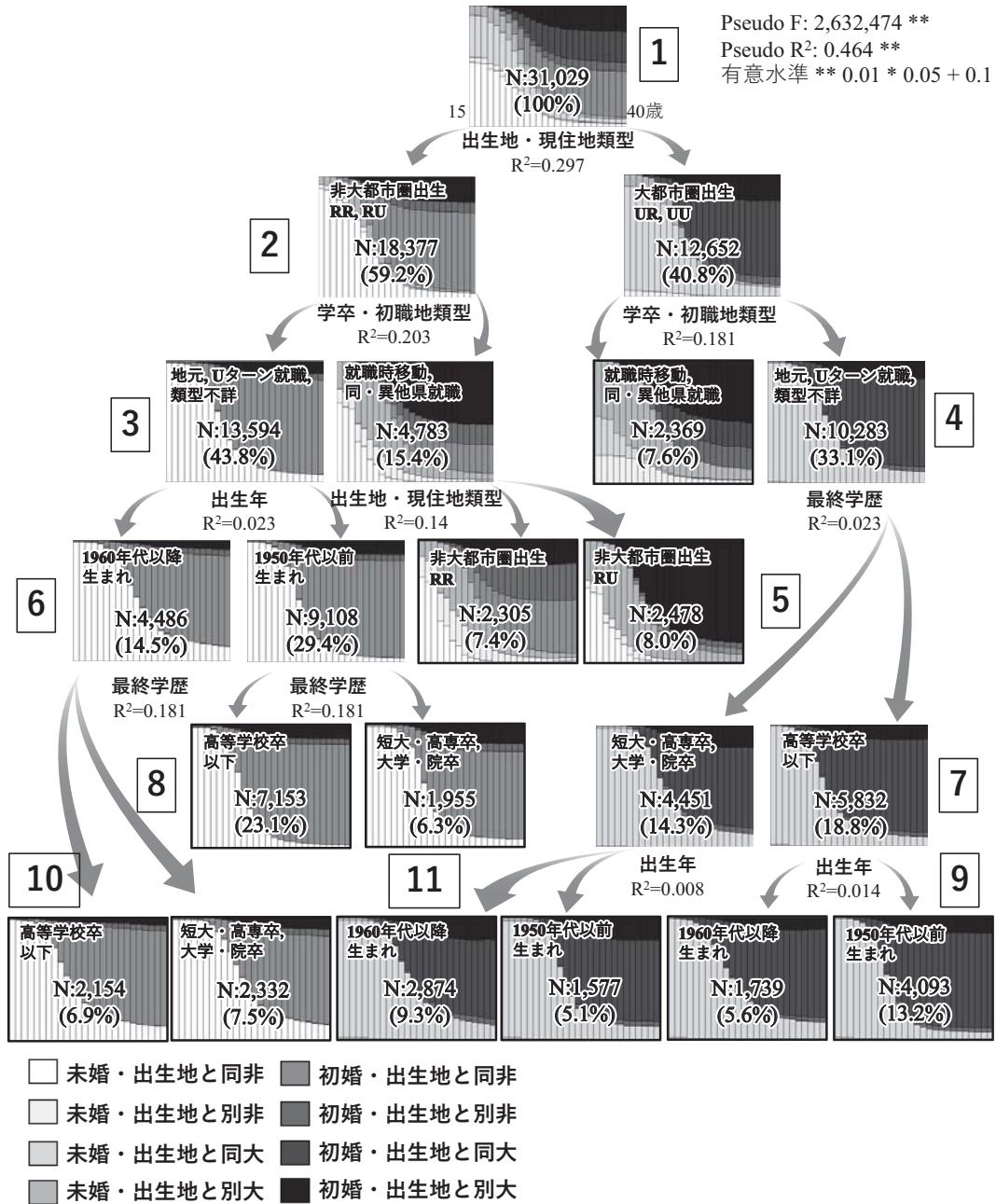
一方で、大都市圏出生の「就職時移動, 同・異他県就職」群は「1950~60年代生まれ」と「1940年代以前生まれ・1970年代生まれ」に分類され (第13階層), 前者は20代前半に移動が生じ, 後者は18歳時点に移動が生じる。非大都市圏出生の方が、就学・就職時期に移動する可能性が高いため、大都市圏出生よりも多様性が高いことがわかる。

(2) 女性の移動経歴・初婚発生モデル

女性の移動経歴・初婚発生モデルについても同様に回帰分析を行った結果、11階層の分類がなされた (図6)。女性モデルの疑似寄与率 (R^2 値) は0.464であった。全体的な傾向は基本的に男性モデルの分析結果と共通する点が多い。第1階層から第4階層までの分類の結果は男性と同様に、出生地・現住地類型および学卒・就職地類型によって、出生地と同じ地域で居住する群とそうでない群に分かれる (第2~4階層)。第2階層は「非大都市圏出生」(RR, RU, 59.2%), 「大都市圏出生」(UR, UU, 40.8%), 第3~4階層は「地元就職, Uターン就職, 類型不詳」群 (非大都市圏43.8%, 大都市圏33.1%) と「就職時移動, 同・異他県就職」群 (非大都市圏15.4%, 大都市圏7.6%) となり、男性に比べ女性は出生地と同じ地域での就職型の構成割合が高い傾向にある。

非大都市圏出生の「地元就職, Uターン就職, 類型不詳」群は出生年によって「1960年代以降生まれ」(14.5%) と「1950年代以前生まれ」(29.4%) に分類される (第6階層)。「1960年代以降生まれ」の方が晩婚・未婚化傾向を示すが、出生地と別の大都市圏居住経験割合は「1950年代以前生まれ」の方が若干高い。さらに「1950年代以前生まれ」は最終学歴によって「高等学校卒以下」(23.1%) と「短大・高専卒, 大学・院卒」(6.3%) に分類され (第8階層), 40歳時点での未婚者割合は後者の方が高い。一方で、「1960年代以降生まれ」も最終学歴によって「高等学校卒以下」(6.9%) と「短大・高専卒, 大学・院卒」(7.5%) に分類され (第10階層), 後者の方が出生地と異なる大都市圏での居住割合が

図6 回帰木分析による、女性の移動経歴・初婚発生モデルの推定結果



高いほか、晩婚傾向を示す。女性は高学歴ほど40歳時点での未婚者割合が高い傾向となる。

非大都市圏出生の「就職時移動，同・異他県就職」群は，現在の居住地が「非大都市圏（RR）」（7.4%）と「大都市圏（RU）」（8.0%）に分類され（第5階層），前者は出生地とは別の非大都市圏で居住する割合が比較的高いものに対して，後者は大半が学卒・就職後に大都市圏へ移動し居住している点に違いがある。

大都市圏出生の「地元就職，Uターン就職，類型不詳」群は，最終学歴によって「短大・高専卒，大学・院卒」（14.3%）と「高等学校卒以下」（18.8%）に分類される（第7階層）。前者は晩婚・未婚化傾向が高く，出生地と別の大都市圏居住傾向も高い。「短大・高専卒，大学・院卒」はさらに出生年によって「1960年代以降生まれ」（9.3%）と「1950年代以前生まれ」（5.1%）に分類され，「高等学校卒以下」も出生年によって「1960年代以降生まれ」（5.6%）と「1950年代以前生まれ」（13.2%）に分けられる。女性は1960年代生まれ以降，高学歴化に伴う大都市圏への移動の増加に伴い，晩婚・未婚化傾向が進んでいる。

（3）移動経歴・初婚発生モデルによる類型化と男女差

図7は男女別に回帰木分析による移動経歴・初婚発生モデルの最終的な類型結果とケース順に示している。図中の両端丸線は男女のタイプの対応関係を示している。以下では，最終的な類型結果による上位5位までの類型について具体的にみていきたい。

男女ともに最も多い類型は，「非大都市圏出生・非大都市圏居住・地元／Uターン就職型」であり，男性は「1940年代生まれ」（14.9%），女性は「高等学校以下で1950年代以前生まれ」（23.1%）の類型である。両類型は40歳時点の未婚者割合が低い「皆婚」世代である点に特徴がある。2位の類型も男女で共通し「大都市圏出生・大都市圏居住・地元／Uターン就職型」で男性は「1940年代以前生まれ」（9.9%），女性は「1950年代以前生まれの高等学校以下の学歴」（13.2%）の類型であり，40歳時点での未婚者割合も低く1位と同様に「皆婚」世代といえる。

3位は男女ともに晩婚・未婚化傾向を示す「少子化」世代である。男性は「大都市圏出生・大都市圏居住・地元／Uターン就職型・短大・高専卒以下・1950年代生」（9.1%），女性は「大都市圏出生・大都市圏・地元／Uターン就職型・短大・高専卒／大学・院卒・1960年代以降生まれ」（9.3%）の類型である。大都市圏に生まれ，大都市圏に居住しているという点では男女に共通点があるが，最終学歴についてみると男性は比較的低い学歴であるほど，女性は比較的高い学歴であるほど未婚化傾向が高いという対照的な傾向を示している。男性3位に対応する女性の類型は女性10位「大都市圏出生・大都市圏居住・地元／Uターン就職型・高等学校卒以下・1960年代以降生まれ」（5.6%）の類型である。一方で女性3位の類型に対応する男性の類型は6位「大都市圏出生・大都市圏居住・地元／Uターン就職型・1950年代生まれ」（6.2%）や9位「大都市圏出生・大都市圏居住・地元／Uターン就職型・大学・院卒・1960年代以降生まれ」（5.6%）の類型である。

4位の類型では，男性は「非大都市圏出生・非大都市圏居住・地元／Uターン就職型・

図7 男女別にみた回帰木分析による移動経歴・初婚発生の類型結果



1950年代生まれ」(8.9%)、女性は「非大都市圏出生・大都市圏居住・就職時移動／同・異他県就職型」(8.0%)の類型である。男性4位に対応する類型は女性6位「非大都市圏出生・非大都市圏居住・地元／Uターン就職型・短大・高専卒、大学・院卒・1960年代以降生まれ」(7.5%)である。一方、女性4位に対応する類型は男性10位「非大都市圏出生・大都市圏居住・就職時移動／同・異他県就職型・東日本／中部出身」(5.4%)および11位「(同上)西日本出身」(5.3%)である。

5位の類型では、男性は「非大都市圏出生・非大都市圏居住・地元／Uターン就職型・高等学校卒以下・1960年代以降生まれ」(7.1%)、女性は「大都市圏出生・大都市圏居住・就職時移動／同・異他県就職型」(7.6%)である。男性5位に対応する類型は女性8位「非大都市圏出生・非大都市圏居住・地元／Uターン就職型・高等学校以下・1960年代以降生まれ」であり、女性5位に対応する類型は男性12位「大都市圏出生・大都市圏居住・就職時移動／同・異他県就職型・1940年代以前／1970年代生まれ」(5.3%)や14位「大都市圏出生・大都市圏居住・就職時移動／同・異他県就職型・1950-60年代生まれ」(5.1%)となる。

全体的な傾向をみると、上位1・2位は男女ともに、大都市圏かどうかにかかわらず、1940～50年代生まれ世代の地元就職・Uターン就職型の「皆婚」世代の類型である。上位3位以降は男女で傾向が異なる。男性は地元就職・Uターン就職型で比較的学歴が低く未婚率が高い類型が上位に位置する。一方で、女性は比較的学歴が高く未婚率も高い、就職時移動、同・異他県就職型の移動類型が上位に位置しているといった違いがみられた。

IV. 考察とまとめ

本稿は「第8回人口移動調査」のデータを用いて、人口移動と家族形成行動の関係性を系列分析により視覚化・類型化することを目的として、移動経歴と初婚発生について分析を行った。

状態分布図・多様性指標によるライフコースの視覚化分析では以下のような知見が得られた。移動経歴については、1) 出生地が非大都市圏か大都市圏かによって移動経歴が異なり、非大都市圏出身者ほど移動経歴の多様性が高い。2) 出生年が近年になるほど（特に1970年代生まれ）、大都市圏出生者の構成割合が高くなるため、年齢別にみた移動経歴の多様性が増加している。一方で、個人のライフコースとしては、移動機会が少なくなるため多様性は減少する。3) 移動経歴は最終学歴の構成と密接に関連し、高学歴者の増加は大都市圏居住者の増加と関連している。高学歴者ほど移動を伴う可能性が高いため多様性が高くなる。4) 「就職時移動」類型は、高等学校卒業後に大都市圏へ移動するが、その後、非大都市圏への移動が生じるため多様性が高くなる。5) 「同他県就職」及び「異他県就職」類型は大学・院卒の移動類型であり、就職した地域において安定的に居住するため、結果として多様性は比較的低いことなどが明らかとなった。

初婚発生については、1) 出生年が近年になるほど、未婚化が進展しているため、年齢

別にみた多様性は増加している。2) 男女ともに最終学歴が高学歴になるにつれて晩婚傾向を示すが、男性では大学・院卒の未婚率は高等学校等に比べて高くなく、女性では高学歴になるにつれて未婚傾向がみられる。その結果、個人のライフコースの多様性という観点からみると、男性は高学歴層ほど多様性が高いが、女性では比較的低くなっている。3) 学卒・初職地類型は学歴構成と密接に関連し、男性では「同他県就職」「異他県就職」＝「地元就職」>「Uターン就職」>「就職時移動」の順で40歳時点の未婚者割合が高く、女性では「同他県就職」「異他県就職」>「Uターン就職」>「地元就職」>「就職時移動」の順となる。

移動経歴・初婚発生モデルについては、最適マッチング分析によるライフコース間の距離行列の算出および回帰木分析による類型化を行った。回帰木分析による類型化は、はじめに「出生地・現住地類型」「学卒・初職地類型」によって大分類がなされ、以降は出生年、最終学歴、出生地域によって分類される。男女ともに1950～60年代生まれを境に、大都市圏出生者割合の上昇、高学歴化による移動類型の変化、晩婚化、未婚化が進んでいる。回帰木分析による最終的な類型の男女差については、上位2位は1940～50年以前生まれの「皆婚」世代の移動経歴・初婚発生パターンであったが、3位以降は男女で差がみられ、男性は地元就職・Uターン就職型で比較的学歴が低く未婚率が高い類型が上位に位置する一方で、女性では比較的学歴が高く40歳時点の未婚者割合も高い、就職時移動、同・異他県就職型の移動類型が類型の上位に位置するといった違いがみられた。

以上みてきたように、系列分析の手法を用いることにより、移動経歴と初婚発生には属性によってライフコースの多様性が異なることを視覚的に明らかにすることが可能である。少子化を牽引する世代である1960～70年代生まれは、それ以前の世代に比べて、親世代の大都市圏への移動の影響を受けて大都市圏出身者の割合が比較的高く、かつ晩婚化、未婚化が進展していることから、年齢別にみたライフコースの多様化は進んでいる。高学歴化は非大都市圏から大都市圏への移動を伴うためライフコースの多様化を促し、また、安定かつ高収入の職に就く可能性が高いため、男性では40歳時点での未婚化を抑制するものの、女性ではむしろ未婚化が進んでいる。このように、ライフコース全体を一つのデータとして分析する系列分析は、これまでの分析では十分に明らかにされてこなかったライフコースの多様性という観点から視覚化、類型化を行うことができる分析手法として社会科学においても有用であろう。

最後に本分析の限界と課題について述べたい。分析の限界については、第1に、系列分析はライフコース全体を時間経過に伴う状態変化として視覚化することが主目的の記述的分析であるため、今回の分析では人口移動と初婚発生の因果関係を必ずしも明らかにしてはいない。因果関係の分析には、イベントヒストリー分析や統計的因果推論などの精密な分析が求められる。第2に、変数選択に関する点として、今回の分析では都道府県間移動と初婚発生との関係のみをみてきたが、都道府県間移動が家族形成行動を説明するために十分かは定かではない。市町村間移動との関係や出生行動との関係についても検討が必要である。第3に、今回は15歳から40歳までのライフコースを観察するという点で40歳以

上のケースに限定したため、主に1940～50年代以前生まれの「皆婚」世代と1960～70年代生まれの「少子化」世代との比較という形での分析となった。少子化の進展に対する人口移動と初婚行動の観点から分析するためには、1980年代以降生まれの近年の動向も含めた分析を行う必要がある。

今後の課題として、「第8回人口移動調査」では、世帯主・世帯員の子ども数を把握することができるため、人口移動経験が子ども数に及ぼす影響について分析が可能である（例えば、小池 2006; 2009; 2014）。また、就学・就職に伴う人口移動が初婚タイミングに及ぼす影響や結婚時の人口移動が地域人口構成や地域の出生力に及ぼす影響についても分析を行うことで、東京一極集中の是正が地域の出生力上昇にどの程度寄与するのか、日本全体の出生力はどのように変化するのか、といった政策課題に対する科学的な根拠を示していくことが必要であろう。

(2019年7月10日査読終了)

付記

本研究は、国立社会保障・人口問題研究所「第8回人口移動調査」の調査票情報を統計法第32条に基づき二次利用した結果を用いている。調査票情報の集計・分析は鎌田が行った。また、JSPS 科研費基盤研究 (C)「人口移動が結婚・出生に及ぼす影響に関する地理学的研究」(研究代表者：山内昌和、課題番号：JSPS17K01241)の助成を受けた。

文献

- 香川めい (2010)「初期キャリアの類型とその規定要因」, 労働政策研究・研修機構『非正規 社員のキャリア形成—能力開発と正社員転換の実態—』労働政策研究報告書 No.117, pp.11-39
- 小池司朗 (2006)「出生行動に対する人口移動の影響について—人口移動は出生率を低下させるか?—」『人口問題研究』62巻4号, pp.3-19.
- 小池司朗 (2009)「人口移動と出生行動の関係について—初婚前における大都市圏への移動者を中心として—」『人口問題研究』65-3, pp.3-20.
- 小池司朗 (2014)「人口移動が出生力に及ぼす影響に関する仮説の検証—「第7回人口移動調査」データを用いて—」『人口問題研究』70巻1号, pp.21-43.
- 国立社会保障・人口問題研究所 (2018)「2016年社会保障・人口問題基本調査 第8回人口移動調査」報告書, 調査研究報告資料第36号.
- 林玲子 (2015)「女性の活躍と人口移動」, 日本学術会議社会学委員会社会変動と若者問題分科会公開シンポジウム「移動する若者/移動しない若者—実態と問題を掘り下げる—」.
- 福田巨孝 (2006)「ライフ・コースは多様化しているか? 最適マッチング法によるライフ・コース分析」西野理子, 稲葉昭英, 嶋崎尚子編『第2回家族についての全国調査 (NFRJ03) 第2次報告書 No. 1: 夫婦, 世帯, ライフコース』, 日本家族社会学会 全国家族調査委員会, pp. 167-181.
- 渡邊勉 (2004)「職歴パターンの分析—最適マッチング分析の可能性—」,『理論と方法』Vol.19, No.2, pp.213-234.
- Aasve, A., Billari, F. C. and R. Piccarreta (2007) "Strings of Adulthood: A Sequence Analysis of Young British Women's Work-Family Trajectories", *Eur J Population* 23, pp. 369-388.
- Abbott A. (1995) "Sequence analysis: New methods for old ideas", *Annual Review of Sociology*, 21: pp. 93-113.
- Abbott A. and J. Forrest (1986) "Optimal matching methods for historical sequences", *Journal of*

- Interdisciplinary History*, 16: pp.471-494.
- Abbott, A. and A. Tsay. (2000) "Sequence Analysis and Optical Matching Methods in Sociology: Review and Prospect" *Sociological Methods & Research*, 29 (1), pp. 3-33.
- Barban, N. and Billari F. C. (2012) "Classifying life course trajectories: A comparison of latent class and sequence analysis", *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 61(5), pp. 765-784.
- Billari F. C. (2001) "Sequence analysis in demographic research", *Canadian Studies in Population*, 28(2): pp. 439-458.
- Billari F. C. and R. Piccarreta (2005) "Analyzing Demographic Life Courses through Sequence Analysis", *Mathematical Population Studies*, 12:2, pp. 81-106.
- Billari, F. C. (2005) "Life course analysis: Two (complementary) cultures? Some reflections with examples from the analysis of the transition to adulthood" *Advances in Life Course Research* 10, pp. 261-281.
- Breiman, Leo, Jerome H. Friedman, R. A. Olshen, and C. J. Stone. (1984) *Classification And Regression Trees*. New York: Chapman and Hall.
- Elder G. H., M. K. Johnson, and R. Crosnoe. (2003) "The emergence and development of life course theory", *Handbook of the Life Course*, pp. 3-19.
- Fulda, B. E. (2016) "The diversity in longitudinal partnership trajectories during the transition to adulthood: How is it related to individual characteristics and regional living conditions?", *DEMOGRAPHIC RESEARCH*, VOLUME 35, ARTICLE 37, pp.1101-1134.
- Gabadinho, A., G. Ritschard, M. Studer and N. S. Müller (2009) *Mining sequence data in R with the TraMineR package: A user's guide*, University of Geneva, (<http://mephisto.unige.ch/traminer>)
- Gabadinho, A., Ritschard, G., Müller, N.S. & Studer, M. (2011), Analyzing and visualizing state sequences in R with TraMineR, *Journal of Statistical Software*. Vol. 40(4), pp. 1-37.
- Gabadinho A., G. Ritschard and M. Studer (2013) Workshop on sequence analysis, New York, October 11th.
- Shannon, C. E. "A Mathematical Theory of Communication", *The Bell System Technical Journal*, Vol. 27, pp. 379-423, 623-656, July, October, 1948.
- Studer, M., Ritschard, G., Gabadinho, A. & Müller, N.S. (2011), "Discrepancy Analysis of State Sequences", *Sociological Methods and Research*. Vol. 40(3), pp. 471-510.

Life-course Analyses of Migration and First Marriage -Classification of Life-course Patterns using Sequence Analysis (Optimal Matching Analysis and Regression Tree Analysis)-

Kenji KAMATA, Shiro KOIKE and Masakazu YAMAUCHI

This paper analyzes the migration types and the occurrence of first marriage for the purpose of visualizing and classifying the relationship between population migration and family formation, by using the data of "The Eighth National Survey on Migration". The sequence analysis is characterized in that the whole life course is regarded as one data, and the analysis was performed using the index which quantified the state change in the life course as "diversity".

As a result, born in the 1960s and 70s, the generation that drives the declining birthrate, the proportion of people born in the metropolitan areas is relatively high compared to the previous generations, the late marriage and unmarried rate are increasing. Therefore, diversification of life courses by age is progressing.

In addition, as a result of classification by regression tree analysis, the type of the top 2 ranks was the migration history and first marriage occurrence pattern of the "marriage-oriented" generation born before 1940-50 years. Although the difference is seen between male and female after the type of 3rd place, the males have a relatively low educational background and high unmarried rate among the types of "the employment in birth prefecture" and "the U-turn employment". While the females have relatively high educational background and high unmarried rate, migration types of "the migration at first employment" and "the employment in same/other regional block" were located at the upper place.