

図2 回帰分析と潜在クラス分析の概念的な比較

をここに挿入

サンプルの行動や態度(意識)から母集団をグループ分けする方法にはクラスター分析も存在する。クラスター分析は、サンプル間の違い(距離)の測定方法を事前に定義し、計測した距離に基づいてサンプルをグループ(クラスター)にまとめていく手法である。これに対して、潜在クラス分析では距離を定義しない。潜在クラス分析では、観測できない要素(潜在変数)に基づいてサンプルの行動や態度(意識)が分類されると仮定する。典型的には、潜在変数が名義変数であると考え、潜在変数の各カテゴリがクラスを示す。また、分類においては、各クラス内において各変数が互いに独立(局所独立)であると仮定して母集団を分類していく。言い換えれば、母集団を局所独立な小グループに分類できる潜在変数を探索していくのが、潜在クラス分析の具体的な計算作業となる(図3)。

図3 母集団を局所独立を基準にグループ分けするイメージ

をここに挿入

母集団を分類する手法としては、変数を取る値の組み合わせでグループ分けする方法((多重)クロス集計の各セルを1つのグループとみなす方法)も考えられる。しかし、多数の組み合わせが存在する場合はこの手法は有用ではない⁷。しかし、潜在クラス分析では数個のクラスに集約することが可能である。

潜在クラス分析では、あるサンプルが特定の(1つの)クラスに属するとは

⁷ 例えば本稿の分析に用いた変数では、就業形態6種類×仕事内容3種類×企業規模6種類=延べ108種類となる。さらに調査年ごとに独立だと考えると、これを調査年の数だけべき乗することになる。

考えず、あるサンプルが各クラスに属する可能性がどの程度ずつあるか、その確率（帰属確率）を求められる点も特徴である。潜在クラス分析は、社会学などの分野で利用されているほか、マーケティングの分野では購買行動から顧客を分類する（セグメントする）方法として、すでに実践的に用いられている。

潜在クラス分析を行う上で留意すべき点は、分析者がクラス数を決める必要がある点である。実際には、クラス数を2、3、4…と増やしながら潜在クラス分析を実行し、各クラス数における情報量規準（AIC や BIC）をもとにクラス数を判断する（モデルを選択する）のが一般的である。クラス数を増やすほどモデルの適合度は向上するが、情報量規準を用いることで、数個の妥当なクラス分けを選択できる。また、情報量規準などの統計量によらず、「理論や仮説をもとにしてクラス数を設定することや、抽出されたクラスが解釈可能かどうかによって最終的なクラス数を決めるという方法もある」⁸。

2. 2. 2 latent transition analysis（潜在移行分析）の概要⁹

潜在移行分析は、基本的にクロスセクション・データを対象とする潜在クラス分析を、パネルデータに応用した手法の1つである。潜在移行分析では、潜在クラスを規定する潜在変数が調査年ごとに存在し、それらが時間の変遷とともに影響を及ぼすと仮定する（図4）。また、観測される変数（顕在変数）と潜在クラス（潜在変数）との関係（各クラスの条件付き確率）が、調査年を通じて同じであると仮定する。これにより、調査年を通じて共通した概念で構成される潜在クラスがどのようなものかやクラスの構成比率の推移、各サンプルが時間の経過とともにどのように推移していくかを見ることができる。なお、前述したように、潜在クラス分析ではあるサンプルが各クラスに帰属する確率が

⁸ 藤原・伊藤・谷岡(2012), p.48。

⁹ 平易な解説としては藤原・伊藤・谷岡(2012)が挙げられる。より詳しい解説としてはCollins and Lanza (2010)が挙げられる。

計算されるため、時間の経過とともにどのように推移していくかは帰属確率をもとにした推移確率として計算される。

図4 latent transition analysis (潜在移行分析) のイメージ (パス図)

をここに挿入

2. 2. 3 本稿の分析方法

本稿ではまず、就業形態、仕事内容、企業規模と調査年とのクロス集計を概観する。これにより、各要素の構成がどのように変化したかをおおまかに確認する。次に、利用するサンプルを調査年を問わずプールして（いわゆる long 形式のデータにして）、潜在変数が1つの潜在クラス分析を行う。この結果を見ることで、潜在クラス分析の結果の見方や手法の特徴を示すとともに、潜在移行分析の結果と比較することで、プールデータとして分析した場合とパネルデータとして利用した場合の結果の違いを見ることが出来る。最後に潜在移行分析を行い、どのような就業状態のグループが抽出され、それがどのように変化していくかを確認する。分析に用いたソフトウェアは、潜在クラス分析には Latent Gold 3.0を、潜在移行分析には Mplus7.0を用いた。Mplus7.0でも潜在クラス分析は実行可能であるが、Latent Gold 3.0の方が推計結果の取扱いや事後的な集計が容易なため、潜在クラス分析には Latent Gold 3.0を使用した。

ただ、あらかじめ述べておくと、今回検討した潜在移行分析のモデルでは計算量が多く実行不可能だったため、潜在移行分析は一部を簡略化して行った。具体的には、すべての調査年のデータを分析に投入すると、クラス数を増やした場合に計算に必要なメモリがMicrosoft Windowsの32bit版で扱えるメモリ領域のサイズを超過するという警告が出て、計算を実行できなかった。そこで、投入する調査年を2年おきに限定して分析した。具体的には、「21世紀成年者縦断

調査」は2003、2005、2007、2009年のデータを、「中高年者縦断調査」は2005、2007、2009年のデータを投入した。これらの調査年を採用した理由は、(1)「21世紀成年者縦断調査」の投入データを2002、2004、2006、2008、2010年の5年分とすると、すべての調査年を投入した場合と同様の警告が出て計算が実行できなかったこと、(2)2004年の労働基準法改正や2006年の高年齢者雇用安定法改正、2008年のリーマンショックやパートタイム労働法改正の影響を見るには、それらの事象が起きた年のデータを投入するよりもその前後の年のデータを比較する方が適していると考えたこと、(3)若年層と高齢層の就業状況の変化に何らかの関連があるかを検討するために、分析に投入する調査年は両調査で揃えた方が望ましいと考えたこと、である。また、調査年を限定してデータを投入して計算は可能になったものの、クラス数を増やすと計算に非常に時間が掛かり、現実問題として計算結果を得るに至らなかったモデルも存在した¹⁰。この点については今後の課題であり、投入するデータの変数の数やカテゴリ数、最適化方法の選択などについて試行錯誤しながら改善を図っていきたい。

3 分析結果

3.1 クロス集計による概観

就業形態、仕事内容、企業規模と調査年とのクロス集計（各変数の構成比率の推移）を図表7¹¹に示した。

表7 クロス集計による概観

ここに挿入

¹⁰ 「21世紀成年者縦断調査」への潜在移行モデルの適用では、2プロセッサ(スレッド)を使用したものの、5クラスモデルでは丸3日掛かって計算が終了しなかった。

¹¹ 表6と表7の違いは、表6では欠損値のサンプルも集計に加えているのに対して、表7では欠損値のサンプルも集計から除いている点である。

まず「21世紀成年者縦断調査」の集計結果を見る。まず、不就業の比率が就業形態では低下し、仕事内容と企業規模では上昇しているが、実数で見ればいずれの項目でも同数である¹²。総合的に考えれば、不就業の割合は2002年から2006年にかけて緩やかな低下傾向にあり、2007年から2010年にかけては緩やかな上昇傾向にあることがうかがわれる。ただし、いずれもわずかな変化である。就業形態を見ると、不就業の動きと呼応するように、正規の割合は2002年から2006年にかけて緩やかな上昇傾向にあり、2007年から2010年にかけては緩やかな低下傾向にある。また、パート（アルバイトとパートの総称）の割合は、不就業と同様に正規と逆の動きになっている。なお、パート（アルバイトとパートの総称）の内訳を見ると、アルバイトが一貫した低下傾向にある一方、パートが一貫した増加傾向にある点が興味深い。このような動きとなる理由は、(1)調査初年度には学生であったサンプルが就職することに伴い、アルバイトが減少したり正規が増加する、(2)調査年度が進む間に、結婚し、結婚後に専業主婦となったりパートとなった、などが仮説として考えられる。なお、今回の分析では「中高年者縦断調査」とカテゴリを揃えるために両者を1つのカテゴリにまとめている。

仕事内容を見ると大きな変化はないが、専門的・技術的な仕事の割合が若干上昇し、販売やサービスの仕事が若干低下する傾向が見られる。前者は就職や熟練による変化、後者はアルバイトが減少することに伴う変化、と推察される。企業規模にも大きな変化はないが、30～299人の割合が低下し、5000人以上・官公庁の割合が若干上昇する傾向が見られた。

次に「中高年者縦断調査」の集計結果を見る。3つの調査項目に共通する不

¹² 前述のとおり、これらの設問の前の設問で「働いていない」と答えたサンプルを「不就業」と再定義しているため、上記の3つの設問いずれでも同数が不就業となる。

就業の動きを見ると、時間の経過とともに割合が上昇する傾向がある。これは、定年退職や、定年退職後に就いた職業をやめることによる影響と考えられる。就業形態を見ると、正規の割合の低下が大きく、ついで派遣・契約の割合の小幅な上昇（特に派遣社員・嘱託）が見られる。これらの傾向は、不就業を除いた比率で見ると、より顕著である。定年退職に伴い、正規から派遣・契約に移動した可能性が考えられる。仕事内容は専門的・技術的な仕事や販売、生産工程・労務作業の比率で小幅な低下が見られる。就業形態と異なり、これらは不就業を除いても大きな割合の変化は見られなかった。定年退職の影響ではなく、世代や時代の影響である可能性がある。企業規模を見ると、不就業以外はいずれの企業規模でも比率の低下傾向が見られるが、不就業を除くと1～4人で比率が増加していることが分かる。定年退職後の再就職先がこの規模の企業になっている可能性がある。

3. 2 プールデータとして潜在クラス分析を実施した結果

前述したように、潜在クラス分析では分析者がクラス数を判断する必要がある。各モデル（クラス数）における情報量規準（AIC と BIC）を表8に示した。「21世紀成年者縦断調査」では BIC が最も小さいのが7クラスモデルで AIC が最も小さいのが9クラスモデル、「中高年者縦断調査」では BIC が最も小さいのが7クラスモデルで AIC が最も小さいのが8クラスモデルであった。ただし、それぞれの最小値付近での AIC と BIC の変化量はわずかであった。

表8 潜在クラス分析：各モデル（クラス数）の統計量
をここに挿入

各クラス数でのクラス分けの結果を示したのが表9である。この図表が示す

内容は次のとおりである。以下、「21世紀成年者縦断調査」について見ていく。まず最上段の「C1/4」は、4クラスモデルにおける第1クラスを示す。2段目は各クラスのサイズ（このクラスに属する確率）である。例えばC1/4の2段目は0.53となっており、全体の53%がこのクラスに属する可能性があることを示している。3段目以下は、各クラスに対する各変数の各カテゴリの応答確率（各クラスにおける各変数の各カテゴリの構成割合）である。C1/4の就業形態を見ると82%が正規、仕事内容は58%がホワイトカラー、企業規模は36%が30～299人、28%が300～4999人となっている。この表の左から3列目には不就業を除いた各変数の構成割合を載せている。これとC1/4の応答確率を見比べると、ホワイトカラーや30～4999人の応答が高いのは全体の構成割合に比較的近い。これらの結果から、C1/4は、正規で中～大企業に勤める確率が高いクラスと解釈することができる。また、C2/4は就業形態、仕事内容、企業規模のいずれも不就業が100%となっている。潜在クラス分析により、適切に不就業が区分けされていることがうかがわれる。

このような作業に基づく解釈結果を最下段に記した。なお、最下段の上に「*」が付してあるクラスは、筆者が、クラス数の増加に伴って新しく発生したクラスと解釈したことを示している。例えば4クラスモデルと5クラスモデルを見比べると、C2/4とC2/5、C3/4とC3/5、C4/4とC5/5は、それぞれ似た応答確率で構成されたクラスとなっており、上から2段目に示されているの各クラスのサイズもほぼ同じである。一方、C4/5は正規が多い点ではC1/4やC1/5と似ているが、仕事内容のホワイトカラーや企業規模の5～29人の応答確率が高い点が他のクラスにない特徴になっている。また、各クラスのサイズを見ると、C1/4が0.53なのに対し、C1/5が0.45、C4/5が0.10で、C1/5とC4/5の合計が0.55となっている。これらの状況から、クラス数を4から5に増加させたことに伴いC1/4からC4/5が分化し、C1/4の後継としてC1/5が存在していると解釈できる。

このような解釈作業を進めると、6クラスモデルでは5クラスモデルの正規のクラス(C1/5)から非ホワイトカラーの正規で中～大企業勤務のクラス(C1/6)が、BICが最小となる7クラスモデルでは5000人以上・官公庁に勤める正規・ホワイトカラーのクラス(C5/7)などが分化している。ただ、7クラスモデルでは7クラス目(C7/7)のサイズが0.02と小さめになっている。

一方、「中高年者縦断調査」のクラス分けの過程を見ると、「21世紀成年者縦断調査」と比べて正規とホワイトカラーがセットとなっている印象を受ける。例えば4クラスモデルでは「21世紀成年者縦断調査」と近いクラス分けになっているが、正規の応答確率が高いC1/4でホワイトカラーの応答確率が0.71と高くなっている。5クラスモデルでは、4クラスモデルの正規・ホワイトカラーのクラス(C1/4)が小～中企業のクラス(C5/5)と中～大企業のクラス(C1/5)に分化する。これは、中高年層では管理職が多いためと推察される。また、その裏返しとして、パートと非ホワイトカラーの結びつきもある程度存在するようだ。6クラスモデルでは、5クラスモデルのパート・非ホワイトカラーのクラス(C2/5)が小～中企業のクラス(C5/6)と中～大企業のクラス(C3/6)に分化する。7クラスモデルでは、5000人以上・官公庁に特化した大企業・正規のクラス(C7/7)が分化するが、サイズが小さめ(0.02)となっている。

「21世紀成年者縦断調査」と「中高年者縦断調査」のクラス分けは別個に行ったものであるが、その過程を比較すると、4クラスモデルでは不就業、自営、正規、パートに分かれる点、5クラスモデルでは正規が企業規模で分かれる点、7クラスモデルで5000人以上・官公庁勤務が分化した点が両調査の結果で共通しており興味深かった。

表9 潜在クラス分析：各モデル（クラス数）の内容

をここに挿入

何クラスのモデルを採用するかについては、「21世紀成年者縦断調査」「中高年者縦断調査」とともに BIC が最小となる 7 クラスモデルが有力な候補である。ただ、7 クラスモデルではサイズが比較的小さい(0.02の)クラスが抽出されることから、6 クラスモデルの方が儉約的との見方もできよう。ここでは、BIC のほか、5000人以上・官公庁勤務が分化した点を考慮し、7 クラスモデルを採用する。

採用した 7 クラスと各属性がどのような関係にあるかを、各属性の分布¹³で示したのが表10である。この図において、各属性の分布を各クラスのサイズ（構成比率）と比較することで、特定の属性が特定のクラスに偏っているかが分かる。なお、これまでは各クラスの表示をサイズが大きい順に記していたが、ここでは「21世紀成年者縦断調査」と「中高年者縦断調査」の比較を容易にするため、両者の似たクラスが対応する順番で表示した（表頭のクラスの記号は表9と同じ）。

「21世紀成年者縦断調査」の結果を見ると、いずれの調査年でも正規・中～大企業のクラス(C1/7)が最も多くなっているが、細かく見ると不就業の比率は2002～2006年にかけて若干の下落傾向にあったものが、2006年以降は若干の上昇傾向にある。年齢別に見ると、23～41歳で正規・中～大企業のクラス(C1/7)が最も多くなっているが、正規・中～大企業のクラス(C1/7)の比率は25～27歳をピークに、加齢に従って少しずつ低下傾向にある。その一方で、非ホワイトカラー・パートのクラス(C3/7)の比率は30代において加齢に伴って上昇する傾向が見られる。また不就業のクラス(C2/7)の比率は30代前半に上昇する傾向が見られる。これらの変動は、結婚や出産を機に、正規・中～大企業のクラス(C1/7)から不就業のクラス(C2/7)、そして非ホワイトカラー・パートのクラス(C3/7)へと推

¹³ 各クラスへの帰属は確率で求められるため、各属性の分布も帰属確率から計算されたものである。

移して女性が存在することを想像させる。性ごとの比率を見ると男女で大きく異なる点も注目される。また、自営のクラス(C6/7)は、調査年ごとに見るとそれほど大きな傾向は見られないが、年齢別に見ると30代前半に比率が大きく上昇する傾向が見られる。

「中高年者縦断調査」では、調査年を追うごとに不就業のクラス(C1/7)の比率が上昇している。年齢別に見ると、50代前半は他のクラスに比べて正規・ホワイトカラー・大企業のクラス(C3/7)の比率低下が目立つ。また60代になると正規・ホワイトカラー・大企業のクラス(C3/7)の比率は一段と低下し、他の就業中のクラスでも60代になると比率が低下する傾向が見られる。その一方で不就業のクラス(C1/7)の比率が上昇している。自営のクラス(C4/7)は50代半ばで比率が上昇する傾向が見られ、62歳まで上昇を続けている。性ごとの比率を見ると、不就業のクラス(C1/7)と正規・ホワイトカラー・大企業のクラス(C3/7)で男女差が大きい。

表10 潜在クラス分析：選択したモデル（クラス数）における属性の分布状況をここに挿入

3. 3 パネルデータとして（潜在推移分析）

潜在移行分析でも、潜在クラスモデルと同様に BIC などを参考に分析者がクラス数を決める必要がある。表11に各モデル（クラス数）における情報量規準（AIC と BIC）を示しているが、クラス数を増やすと計算負荷が高まって計算に数日かかっているため、ここでは「21世紀成年者縦断調査」で4クラス、「中高年者縦断調査」で6クラスまでの提示にとどまっている。「21世紀成年者縦断調査」と「中高年者縦断調査」のいずれでも AIC や BIC はクラス数の増加とともに下がっており、計算した範囲では増加に転じていない。このため、AIC や BIC

から適切なクラス数は判定できていないが、個々のモデル（各クラス数のモデル）の計算は完結しているため、以下では計算できた範囲で各クラスの内容を見ていく。

表11 潜在移行分析：各モデル（クラス数）の統計量

ここに挿入

各クラス数でのクラス分けの結果を示したのが表12である¹⁴。「21世紀成年者縦断調査」の結果を見ると、3クラスモデルにおいて就業者が企業規模が30人以上か未満かでクラスが分かれ、4クラスモデルでは就業者のクラスが30人未満・自営業、ホワイトカラー・正規、非ホワイトカラーに分かれた。プールデータとして潜在クラス分析を実施した結果（図表9）の4クラスモデルと比べると、潜在クラス分析では就業形態を中心に分化していたのに対し、パネルデータとして潜在移行分析を実施した結果では仕事内容や企業規模を中心に分化しているのが興味深い。また、各クラスのサイズは調査年によって大きな違いはなかったが、不就業のクラスは調査年が新しくなるにつれて少しずつ小さくなる傾向が見られた。

「中高年者縦断調査」の結果を見ると、3クラスモデルにおいて就業者が自営かどうかを中心にクラスが分かれ、4クラスモデルでは勤め人が正規かパートかで分かっている。5クラスモデルでは正規が企業規模によって分かれ、6クラスでは自営が仕事内容で分かっている。プールデータとして潜在クラス分析を実施した結果（図表9）と比べると、4クラスモデルや5クラスモデルは

¹⁴ 潜在移行分析では図4で示したように調査年ごとに潜在変数（潜在クラス）を推計しているが、観測される変数（顕在変数）と潜在クラス（潜在変数）との関係（各クラスの条件付き確率）が調査年を通じて同じであるという仮定を組み込んで推計するため、表12の形式で示すことができる。条件付き確率は調査年を通じて同じだが、クラスのサイズは投入した調査年ごとに異なる。

潜在クラス分析の結果と潜在移行分析の結果は比較的似ている。6クラスモデルは、潜在クラス分析ではパートが企業規模を中心に分化したのに対し、潜在移行分析では自営が仕事内容で分かれたのが興味深い。なお潜在クラスモデルでも、8クラスモデルにおいて自営が仕事内容で分化している。

表12 潜在移行分析：各モデル（クラス数）の内容

ここに挿入

次に、潜在移行分析の特徴である調査年間の推移状況を見ていく。表13は、2時点間の推移確率を示したものである。例えば「21世紀成年者縦断調査」の2クラスモデルの「2003\2005」の欄は、2003年のC1クラス(2クラスモデルなので、表頭にはC1/2と表記)は、2005年には67%の確率でC1クラスへ移行し、33%の確率でC2クラスへ移行することを示している。なお、ここで示している推移確率は、各サンプルの各クラスへの帰属確率をもとに計算されたものである。以下では「所属クラス」等の表現を用いているが、厳密には、あるサンプルが特定の1つのクラスに属する形で計算されたものではない点には留意されたい。

「21世紀成年者縦断調査」の結果を見ると、各2時点間ともに、クラスが変わらない確率が各クラスの中で最も高い。4クラスモデルを見ると、2007年から2009年への移行でクラスが変わらない確率が特に高く、不就業以外の3クラスでは90%を越えている。この理由は定かではないが、この2時点間の間にリーマンショックが発生したことが影響している可能性もある。このことは、今後の調査年のデータを追加したり、当分析には投入していない退職事由の設問などを分析することなどで確認する必要がある。

一方、クラスが変わる確率を見ると、いずれの2時点間でも不就業のクラス

(C2/4)から非ホワイトカラーのクラス(C4/4)へ移行する確率が高い。また、2003年から2005年と2005年から2007年では非ホワイトカラーのクラス(C4/4)からホワイトカラー・正規のクラス(C1/4)へ移行する確率が11%あったが、2007年から2009年ではゼロ%になっている。同様にホワイトカラー・正規のクラス(C1/4)へ移行する確率をみると、30人未満・自営のクラス(C3/4)からの移行は、2003年から2005年と2005年から2007年では2~3%あったが2007年から2009年ではゼロ%になっている。また、2007年から2009年では、不就業以外の3クラスから不就業のクラス(C2/4)へ移行する確率が各クラスからの移行確率の第2位となっており、その確率は第3位以下と大きな差がある。このような2007年から2009年への移行確率に特有の傾向にリーマンショックの影響があるかどうかは、追加的な分析の対象となる。

「中高年者縦断調査」の結果を見ると、ここでも各2時点間ともに、クラスが変わらない確率が各クラスの中で最も高い。ただし、「21世紀成年者縦断調査」でみられた2007年から2009年への移行でクラスが変わらない確率が特に高いという傾向は、このデータでは必ずしも見られず、多くのモデルで不就業のクラス(C3/4)にとどまる確率のみが2007年から2009年への移行で少し高くなっていた(2005年から2007年への移行では0.843、2007年から2009年への移行では0.879)。

「中高年者縦断調査」の結果の特徴としては、いずれのモデルのいずれの移行タイミングでも、自営のクラス(C2/3、C2/4、C4/5、C3/6、C5/6)にとどまる確率が、他のクラスにとどまる確率よりも高かった。これは、自営以外の就業者のクラスでは定年退職によって職にとどまれないケースがある一方で、自営では自身の意志で引退時期を決められる影響かもしれない。

クラスが変わる確率を見ると、不就業以外のクラスから不就業へ移行する確率が各クラスからの移行確率の第2位となっており、その確率は第3位以下と大きな差がある点が挙げられる。この傾向は「21世紀成年者縦断調査」と同様

である。また、不就業のクラスではパート・非ホワイトカラーのクラスに移る確率が高い点も、「21世紀成年者縦断調査」と同様の傾向であった。

表13 潜在移行分析：各モデル（クラス数）の2時点間の推移確率
をここに挿入

次に、分析対象期間全体の推移パターンの確率（比率）をみる。2時点間の推移確率から想像されるとおり、分析対象期間を通じて所属クラスが変わらない確率が高かった。「21世紀成年者縦断調査」と「中高年者縦断調査」を比べると、「中高年者縦断調査」の方が所属クラスが変わらない確率が高かった。

所属クラスが変わるパターンのうち確率が高いものを見ると、「21世紀成年者縦断調査」の4クラスモデルでは不就業のクラス(C2/4)か非ホワイトカラーのクラス(C4/4)を経てホワイトカラー・正規のクラス(C1/4)に至るパターンと、不就業のクラス(C2/4)を経て非ホワイトカラーのクラス(C4/4)に至るパターンの確率が高めであった。なお、これらの確率が高めのパターン（上位5つ）では、2007年と2009年の所属クラスが変わらない点が共通していた。「中高年者縦断調査」では、2009年に不就業のクラスに所属するパターンの確率が高めであった。

表14 潜在移行分析：各モデル（クラス数）の分析対象期間全体の推移パターンの確率
をここに挿入

4 結論と今後の課題

本稿では、厚生労働省が実施した「21世紀成年者縦断調査」と「中高年者縦断調査」という大規模パネルデータに潜在移行分析(latent transition analysis)を適

用して、2000年代の若年層と高齢層の就業状態の変化パターンを分析した。計算環境の制約により十分には分析できなかったが、現時点では、(1)「21世紀成年者縦断調査」において、2007年から2009年にかけては、2003年から2005年や2005から2007年と比べて就業状態（クラス）の変化が少なく、特に非ホワイトカラーのクラスや自営のクラスからホワイトカラー・正規のクラスへ移行する確率はゼロとなった、(2)「21世紀成年者縦断調査」では所属クラスが変わるパターンのうち不就業のクラスか非ホワイトカラーのクラスを経てホワイトカラー・正規のクラスに至るパターンの確率が高めだった、(3)「中高年者縦断調査」では自営のクラスにとどまる確率が他のクラスにとどまる確率よりも高かった、という結果を得られた。また、潜在移行分析の概要や特徴を示すとともに、大規模パネルデータに適用する際に生じる計算負荷について示すことができた。

これらの結果を受けて、今後の課題として、計算方法の工夫を通じたこの分析手法の改善や拡張、別のアプローチによる分析の深耕の必要性が明らかになった。可能であれば、毎年の調査データの投入や投入変数の追加やカテゴリ化の改善、生まれ年ごとの分析や事後集計によって、問題の所在をより細かく明らかにしたい。そのためには計算方法の工夫が必要だが、本格的な改善は専門家にゆだねざるを得ない。統計パッケージのユーザーとしては、パッケージで用意された設定の試行錯誤や投入データの工夫、統計パッケージの比較などによって対処していきたい。今回、潜在移行分析による探索的な分析により興味深い問題の所在は明らかになったが、要因分析にあたっては検証的な分析など他のアプローチによる深耕が必要である。特に、「21世紀成年者縦断調査」において、2007年から2009年にかけては就業状態（クラス）の変化が少ない点については、分析対象期間の追加など、新たな分析が求められる。また、今回は就業状態を分析対象としたが、心理的な態度や健康状態などこの調査で得られる他の情報に対する潜在移行分析の適用についても、検討していきたい。

補論 データの整備について¹⁵

今後、「21世紀成年者縦断調査」や「中高年者縦断調査」を分析する際に参考となる可能性があるため、両データのデータ整備について下記のとおりまとめる。

補論. 1 「21世紀成年者縦断調査」における同一個人の識別

パネルデータとして利用するためには、複数の調査年をまたいで同一個人を識別することが必要になる。しかし、厚生労働省から提供された「21世紀成年者縦断調査」のデータは調査年ごとに別のファイルとなっており、個人を明確に識別するための継続的なコードが付与されていない。

理屈の上では「地区番号」「単位区番号」「世帯番号」「該当者番号」を組み合わせれば個人を識別できるはずである。筆者らは統計ソフト STATA の Group 関数を用いて「地区番号」「単位区番号」「世帯番号」「該当者番号」の組み合わせごとに新たな ID コードを付与したが、同一の ID コードで調査年によって生まれ年や生まれ月が変わっているサンプルがあった。これらの中には、生まれ年が1年だけ違うなど単純な回答ミスと思われるものもあったが、まったく違うものも存在した。同一の ID コードで調査年によって生まれ年や生まれ月が変わっているサンプルは同一人物ではない可能性があるため、相違の程度を問わず一律に分析から除外した。除外したサンプルの数は、いわゆる long 形式のデータで、女性が97163サンプル中577サンプル、男性が100587サンプル中767サンプルであった。

同一の ID コードで調査年によって生まれ年や生まれ月が変わる理由は明らか

¹⁵ これらの処理を含む基礎的なデータ整備は北村智紀氏(ニッセイ基礎研究所)の労によるものである。記して謝す。

ではないが、例えば、離婚したあとに別人と再婚した場合に、新旧の配偶者は別人であるにも関わらず新旧の配偶者に同一の該当者番号が付与される可能性があるのではないかと考えている。この仮説が正しければ、該当者番号と生まれ年と生まれ月が同一だからといって毎年の配偶者が同一人物でない可能性も考えられるが、実態の確認はできないためこの可能性は無視した。

補論. 2 本稿における「21世紀成年者縦断調査」の本人/配偶者・男/女データの取扱い

厚生労働省から提供された「21世紀成年者縦断調査」のデータは、該当者番号を男性と女性で別のデータ項目に格納し、(設問内容が同じであっても)男性用の調査票の回答と女性用の調査票の回答を別のデータ項目に格納した上で、1夫婦もしくは1単身者を1つのレコードしている。そこで、まず、男性の該当者番号が存在するレコードと女性の該当者番号が存在するレコードをそれぞれ抽出して、男性と女性のデータセットをそれぞれ作成した。次に、データ項目(変数)を、分析に使う男女に共通したもののみ限定した。その上で、レコードを追加する形で男女のデータセットを(縦に)結合した。

この処理により分析対象のサンプル数は増えたが、(1)夫婦としての状態(2人の状態の組み合わせ)を考慮していない、(2)配偶者の年齢は多様であるため2002年10月末時点の20~34歳以外の対象者も含むサンプルを含む、という状況が発生している点には留意が必要である。

なお、2003年調査については単身男性の回答が女性のデータ項目に格納されているため、上記の結合処理を行う前にこれらを男性のデータ項目に移動した。なお、本稿のように男女のデータを結合する場合にはこの処理を行う必要はないが、当研究プロジェクトの他の研究では必要だったため行った。

参考文献

- Collins, Linda M. and Stephanie T. Lanza (2010) *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- 佐藤一磨(2010)「景気後退期の就業行動の変化」瀬古美喜・照山博司・山本勲・樋口美雄編『日本の家計行動のダイナミズム VI』慶応義塾大学出版会, pp.89--113.
- 四方理人(2008)「有期労働契約期間の規制緩和と雇用の安定」, 樋口美雄・瀬古美喜・慶応義塾大学経商連携21世紀COE編『日本の家計行動のダイナミズム IV』慶応義塾大学出版会, pp.141--159.
- 武内智彦(2013)「JGSS統計セミナー2012: パネルデータを用いた潜在クラスモデル分析」大阪商業大学JGSS研究センター編『日本版総合的社会調査共同研究拠点 研究論文集 [13]』(http://jgss.daishodai.ac.jp/research/res_top.html) .
- 尾崎幸兼(2007)「潜在推移分析」豊田秀樹編『共分散構造分析 [Amos 編]: 構造方程式モデリング』東京図書, pp.190--191.
- 野村総合研究所(2012)『日本におけるパネルデータの整備に関する調査 報告書』(平成23年度内閣府大臣官房統計委員会担当室請負調査) (http://www5.cao.go.jp/statistics/nenpou/nenpou_1206.html#h6) .
- 藤原翔・伊藤理史・谷岡謙(2012)「潜在クラス分析を用いた計量社会的アプローチ: 地位の非一貫性、格差意識、権威主義的伝統主義を例に」『年報人間科学』(33), pp.43-68, (<http://hdl.handle.net/11094/12447>)
- 松田紀之(1988)「潜在クラス分析への応用」『質的情報の多変量解析』朝倉書店, pp.166--197 (<http://www.sci.kagoshima-u.ac.jp/~ebsa/matsuda01/index.html>) .
- 山口和範(2004)「潜在クラス分析」『図解入門 よくわかる統計解析の基本と仕組み 改訂版』秀和システム, pp.211--220.

- 山本勲(2008)「高年齢者雇用安定法改正の効果分析：60歳台前半の雇用動向」樋口美雄・瀬古美喜・慶応義塾大学経商連携21世紀COE編『日本の家計行動のダイナミズム IV』慶応義塾大学出版会, pp.161--173.
- 渡辺美智子(2001)「因果関係と構造を把握するための統計手法—潜在クラス分析法—」岡太彬訓・木島正明・守口剛編『マーケティングの数理モデル』朝倉書店, pp73-115.

表1 国内パネル調査の概要一覧

調査名 (英語略称)	21世紀 出生児 縦断調査	21世紀 成年者 縦断調査	中高年者 縦断調査	働き方とライフ スタイルの変化 に関する全国 調査(JLPS) [高卒パネル 調査]	働き方とライフ スタイルの変化 に関する全国 調査(JLPS) [若年パネル 調査]	働き方とライフ スタイルの変化 に関する全国 調査(JLPS) [社年パネル 調査]	慶應義塾家計 パネル調査 (KHPS)	日本家計 パネル調査 (JHPS)	消費生活に 関する パネル調査
概要・ 目的	子どもの成長・ 発達の様子 や、子育てに 関する環境や 意識、行動の 変化を把握	成年者の 結婚、出産、 就業等の実態 及び意識の 経年変化の 状況を把握	団塊の世代を 含む中高年者 の健康・就業・ 社会活動につ いて、意識面・ 事実面の変化 の過程を継続 的に把握	雇用環境の変化や、少子高齢化社会の到来と いった社会変化が、人々の働き方、 ライフスタイルに与える影響を把握			社会全体の人口構成を反映し た家計パネル調査の日本国内 での整備		収入・支出・貯 蓄、就業行動、 家族関係など の若年女性の 生活実態の 把握・分析
対象	全国の2001年 (平成13年) 1月10日から 同年7月17日 の間に出生した 子ども(平成13 年)及び2010年 (平成22年)の5 月に出生した子 ども(平成22年 パネル)	2002年(平成 14年)10月末 時点で20～34 歳であった全 国の男女及び その 配偶者	2005年(平成 17年)10月末 時点で50～59 歳である 全国の男女	無業率、進学 率の傾向の違 いを反映させ た、神奈川県、 宮城、石川、秋 田の4県で、20 04年1月から3 月にかけて高校 を卒業した生 徒	日本全国に居 住する20～34 歳の男女	日本全国に居 住する35～40 歳の男女	日本全国の 一般世帯・個 人	日本全国の 一般世帯・個 人	24歳以上の 若年女性 (4コーホート)
実施 時期	2001年より 年2回 (平成22年パ ネルの調査は、 1月生まれは1 月、7月生ま れは7月に実 施)	2002年より 年1回	2005年より 年1回	2004年より 年1回	2007年より 1～3月の期間 で年1回	2007年より 1～3月の期間 で年1回	2004年より 年1回	2009年より 年1回	コーホートA:1993 年～(24～34 歳女性) コーホートB:1997 年～(24～27 歳女性) コーホートC:2003 年～(24～29 歳) コーホートD:2008 年～(24～28 歳)
調査 手法	郵送法	訪問留置法 2010年 より郵送法	訪問留置法 2010年 より郵送法	郵送法	郵送配布、訪 問回収	郵送配布、訪 問回収	訪問留置法	訪問留置法	訪問留置法
回収 数	35,264人 (2010年) (平成13年 パネル)	18,025人 (集計対象数 は 14,755人) (2009年)	26,220人 (2010年)	約500人	約2,700人	約1,400人	約4,000人	約4,000人	約2,100人 (第18回調査 2010年10月 実施時点)
実施 主体	厚生労働省			東京大学 社会科学研究所			慶應義塾大学 パネル調査共同研究拠点		家計経済 研究所
実査 委託先	訪問留置法による調査は地方自治体の 統計調査員が実査			中央調査社			中央調査社		中央調査社

注1)長期にわたる調査のため調査手法が年によって異なる場合がある。
注2)回収数は調査回によって変動するため目安を記載している。