

自殺率に関する要因分析のための変数群識別戦略の提案¹

—自己組織化マップの適用によるパイロットスタディー—

与謝野有紀²

【要約】

自殺は社会的信頼の欠如が助長する社会病理現象の一つである。日本は1998年に自殺者数の急増を経験し、1998年からの14年間にのべ45万人以上が自殺で亡くなっている。その後、自殺者数は2万人台にとどまっていたが、2022年には男性の自殺数が13年ぶりに増加しており、さらには、小中高生の自殺が初めて500人を超えた。このように、自殺は現在においても喫緊の課題といえる。これまで、日本の自殺率の変動の原因を探ろうとする計量分析が、多数行われてきた。しかしながら、1998年の自殺率の急騰の原因に関して、研究者間で一致した見解が得られているとはいえない。また、自殺率の変動に関する決定的な因果モデルも得られていない。本研究は、この研究状況を打破するために、新規の分析戦略を提案するものである。具体的には、因果分析に先立って、自己組織化マップを用い、地域のクラスタリングを行うことで、変数群を効率的に識別しようとするものである。この分析戦略は、8つのステップから成っている。本研究では、このうちの第6ステップまでを実際のデータを用いて実施し、新たな分析戦略の有効性を示した。

キーワード：自殺、自殺率、自己組織化マップ、変数群識別戦略

1. 自殺率の要因分析をめぐる課題

自殺は社会的信頼の欠如が助長する社会病理現象の一つである(Kawachi and Kennedy 2002; Kelly et al. 2009; 与謝野 2011)。日本の自殺者数は1998年に32,863名となり、統計を取り始めて以来、3万人を初めて超えた。前年の1997年から8,000人以上自殺者が増加しており、増加数の約78%が男性であった。その年から2011年までの14年間、日本の自殺率は3万人を超え続け、総計45万人以上が自殺で亡くなっている。また、この期間、自殺は比較的若い世代での死因のトップを占めていた。たとえば、2010年の統計では、男性の場合、15～39歳で自殺が死因のトップであり、女性の場合、15～34歳の死因のトップとなっていた。こうした状況から、自殺対策は喫緊の課題となり、自殺対策基本法が2006年に施行され、自殺総合対策大綱が2007年に策定されるなど、国を挙げての自殺対策が実施されることとなった。また、自殺予防総合対策センターが国立精神・神経医療セ

¹ 本研究は、2020年度関西大学学術研究員研究費によって行った。

² 関西大学社会学部教授

ンター・精神保健研究所内³に設立され、臨床的な対策の実践と並行して、自殺の要因分析が行われてきた⁴。しかしながら、1998年における自殺者数急増の原因ははまだ特定されておらず、また、人口や年齢を調整した自殺率の要因分析も断定的な結論を得ているとはいいがたい⁵。

その後、自殺者数は2万人台にとどまってはいるが、昨年(2022年)には男性の自殺数が13年ぶりに増加している。さらには、小中高生の自殺が初めて500人を超えるなど、自殺は現在も喫緊の課題であり、原因を探る分析手法の改善もせまられている。そこで、本稿は、自殺率の分析に対して、方法論的に新規な分析戦略の提案を目論む。すなわち、既存の因果モデル分析の手続きとは異なり、まず自殺率と社会環境をめぐる地域パターン分類から始め、それを基礎としてパラメトリックなモデル適用をするという戦略の提案である。ここでは、自殺と社会環境をめぐる地域パターン分類のパイロットスタディまでを行い、その有効性を確認する。

1-1. 自殺率の要因分析の視角

まず、自殺率の要因分析が暗黙に想定する理論枠から整理したい。自殺率は、E. デュルケームの古典的研究にみられる通り、個人的行為の指標ではなく、社会的現象の測度と位置付けられてきた(Durkheim 1897=2002)。すなわち、自殺率は、自殺という個人的行為の地域的集計値であり、個人的行為を超えたマクロ的測度として機能することが想定されている。言い換えれば、自殺率の分析は、個別の自殺の原因を探るのではなく、地域間での自殺の生じやすさの差の社会環境要因を探ろうとするものといえる。そのため、自殺という個人的行為の直接の原因を明らかにすることは原理的にできず、個別の自殺およびその原因を促進・抑制するような社会環境要因を明らかにする作業といえる。

例えば、失業率によって、地域間の自殺率の差異がかなりの程度予測できたとしても、自殺という個人的行為に関しては、次のミクロ的分析への出発点を与えたものと位置づけられるにすぎない。実際、失業がどのような形で自殺の起こりやすさにつながっているのかに関しては、経済的困窮仮説、社会的孤立化仮説、生活習慣仮説⁶とも呼ぶべき複数の仮

3 現在は、国立研究開発法人国立精神・神経医療研究センターとなっている。

4 自殺予防総合対策センターと統計数理研究所の連携による自殺数の集計、推計などが積極的におこなわれた(藤田 2009)。また、自殺率のベイズ推定などは、自殺予防対策センター(2011)にまとめられている。

5 たとえば、池田・伊藤(1999)は、団塊の世代が自殺の多い高齢者に入りつつあるためとしているが、残念ながら計量的な根拠は希薄である。また、佐々木(2005)は、自殺率の変化について解析しているが、社会史的なアプローチであり、計量的な根拠を与えていない。また、自殺率の理解にも疑問が残る。Watanabe et.al(2006)は失業率、自己破産率、離婚率に自殺率を回帰しているが、1998年のドラスティックな変化を説明しているとはいいがたい。計量分析含め、自殺率の研究は多数あるが、いずれも、1998年における急激な自殺率の上昇を説明できているとはいいがたい。

6 この呼び方は筆者が、Martikainen and Volkonen(1996)の議論を要約して命名したものである。正確には以下の通り、健康状態の問題などを含む仮説である。Martikainen and Volkonenによると、低失業期に失業していたものは、高失業期に就労していても自殺しやすい傾向にあることを指摘している。すなわち、低失業期に失業していたものは食事、生活習慣、健康状態などでもともと問題が大きく、失業は原因ではなく、自殺を生みやすい生活習慣等の結果であるとの分析結果を示している。ま

説があり、計量的にもいまだ議論のあるところである。こうした研究状況が意味するところは、自殺率の研究は、ミクロな状況の想定とは独立に展開しうるものであり、マクロな社会指標間の連関構造を明らかにする試みとなっているということである⁷。

こうした状況を暗黙の前提として、自殺率のマクロ変数間の連関構造の解明を試みた研究は、医学系、人文・社会科学系の多分野でなされている。そのため、臨床医学的な自殺研究に次いで研究資源が投入されてきた分野とってよいだろう。こうした、積極的かつ真摯な研究の積み重ねにもかかわらず、日本の自殺率の増減あるいは地域間格差の背景にある社会環境変数の同定は、いまだ十分になされているとはいえない。既存研究の多くが、自殺と関連する社会環境変数を識別したとする計量分析結果を導き出しているが、研究結果間の結論の一貫性が高いとはいえず、学術的に合意の取れた説明モデルも導出されていない。このことは、1998年に男性の自殺数が急増し、自殺数が全体として三万人を大きく超えた理由について、衆目の一致する議論がないことから明らかである。

1-2 マクロ変数間の因果モデルをめぐるミクロ的課題

前述のとおり、自殺率の説明モデル（＝マクロ変数間の因果構造）が適切に識別できたとしても、自殺という個人的行為が生じる原因を識別するには別のアプローチが必要になる⁸。たとえば、自殺率と失業率に相関関係がみられる場合、前述のとおり複数の仮説で自殺という個人的行為をめぐる因果が説明されている。たとえば、経済的困窮仮説では「失業により経済的困窮者が増え、生活苦による自殺者が増大する」と仮定されており、社会的孤立化仮説では「失業による役割の喪失が孤立感を高め、自殺にいたる精神状態を出現させやすくすることで自殺者が増大する」と仮定される。

後者のプロセスを実証するためには、「失業者が社会的に排除されているか」、「失業者が役割の喪失感を抱いているか」、「失業者がはく奪感を抱いているか」、「役割の喪失感、はく奪感が自殺にいたる精神状態を出現させるほど深刻なものか」が測定されなければならない。これは個人レベルでのミクロなデータでの検討の必要性を示唆する⁹。また、前者については、一見自明に見えるが、それでも「失業により経済的困窮にいたったものはどのくらい増えたか」、「経済的困窮がどの程度自殺を迫る状況を生み出しているか」が測定されなければ仮説の妥当性は検証できず、やはりミクロなデータの収集の必要性が示唆され

た、その他、スティグマ効果も議論している。

7 この点を、自殺対策との関係で整理するならば、「自殺率の変化の背後にある社会環境変数を制御することで、自殺率をコントロールする」という意図が自殺率の要因分析の背景に存在しているといえるだろう。マクロな社会政策でマクロ現象としての自殺率を制御できるという理論枠が背景にあるものといってもよい。

8 これは、被説明変数が個人的行為の集計である場合の分析、いわゆる生態学的相関を基礎とした分析一般に起こりえる事柄であり、自殺に限定されない。

9 佐々木(2006)は、医学における遺伝子レベルの研究をミクロと位置づけ、個人のデータの集積による研究をマクロ的な研究と呼んでいる。本稿ではこれと異なり、自治体、医療圏などの地理的範囲を単位として与えられているデータをマクロなデータ、個人単位の分析が可能なデータをミクロなデータとして対比している。

る。すなわち、自殺率のマクロな相関分析は、それ自体では個人的行為である自殺を説明するモデルとしては完結しえない。

次項では、ミクロなデータの詳細な分析を行う臨床医学的な手法が、個人的行為としての自殺の要因分析として有効であること、その一方、臨床医学的手法の有効な活用のためには、マクロな自殺率の分析が必須であることを整理する。

1-3 自殺率に関するマクロ分析の必然性

自殺率の研究とは対照的に、自殺対策の重要な柱をなす臨床医学的な研究では、個人単位の検討が可能なアプローチが取られてきている。すなわち、臨床の対象としての個人に重点が置かれ、特定の事例を基礎とした原因の推測と対策が試みられている(河西 2008, 2011; 松本ほか 2013 など)。この成果は、秋田モデルなどにも反映され、有効性を発揮してきている¹⁰。

こうした実践的成果の高さと裏腹に、臨床医学的な研究は、社会状況のマクロな変化がもたらす影響については推測することができないという特質も有している。すなわち、個人的行為としての自殺の要因を変化させるマクロ指標の効果を識別できないという特質である¹¹。また、ミクロな原因変数と特定の少数の対象に集中した検討がなされやすいため、検討対象(特定地域の特定集団)以外の自殺の変動については慮外となってしまう傾向がある。そのため、検討対象以外を含む一般的な自殺率の変動要因の識別は臨床研究からは得られず、結局、マクロな自殺率の変化から推測せざるをえないという状況が乗じる。

先述のとおり、マクロな自殺率の研究は、個人的行為としての自殺の原因を明らかにすることはできない。しかしながら、マクロな自殺率の研究は、自殺率と社会環境の一般的関係を整理し、その知見のもとに自殺対策に資することを目論む研究として機能する。この研究・実践アプローチは、以下のような実践的な知見を与える。

もしも、自殺率とマクロ変数間の連関構造が識別され、自殺率と因果関係をもつ独立変数が識別されたならば、対策は二つの方向で行われうる。第一は、マクロ変数を制御し、自殺率が下がる方向に当該マクロ変数の値を変化させるというものであり、第二は、自殺を助長するミクロな原因変数に、マクロ変数が影響を与えないような対策をするというものである¹²。後者の成功例は、秋田での官学民連携の自殺対策といえよう。そこでは、マクロな社会環境変数の大きな改善を待つ前に、個人レベルにおいて、社会環境の悪影響をできるだけ早く阻止しようとする対策がなされていた。ただし、こうした地域限定の成功モ

10 官民学の連携によって構成される自殺対策モデルであり、最新の「【第2期】秋田県自殺対策計画(素案)」(2023年3月)にも反映されている。

11 ここで「短所」「弱点」などの用語を用いていないのは、臨床的研究の目指すところがここでの問題意識と異なるため、こうした用語の利用が不適切と考えるためである。

12 こうした見方は、紺田・与謝野(2013)が「『ミクロなプロセス』が発動する蓋然性を高める『マクロなプロセス』の対策も十分に必要がある」とした見方から一貫している。

デルは、一般的に言って、全国網羅的に有効性を発揮することは困難だといえる¹³。

この議論は、「マクロ変数がミクロ変数へ与える効果を統制する」という方法と本質に関係する。たとえば、ある個人に対するミクロな臨床的自殺対策モデルが「少子高齢化の進む農村部で、男性の高自殺率の対策として構築され」、かつ「成功した」場合を考えよう。この対策モデルは「少子高齢化の進む農村部で、男性の自殺率が高い地域」では、モデルの発祥地域同様に大きな効果を持つことが期待できる。しかし、大都市でこの対策モデルが大きな効果をもつかどうかは不明である。なぜなら、社会環境変数が大きく異なるからである。

この点を都市規模、失業率、自殺率の分析を例にとって考えてみよう¹⁴。

図 1 は、2005 年の全市区町村の失業率と男性の自殺率¹⁵を、市区町村の規模別にプロットしたものである。大都市圏では、失業率が自殺率と高い相関関係を有しているが、都市規模が小さくなるほどその関係は消失する傾向にあり、郡部ではほぼ無関係となっていることが分かる。

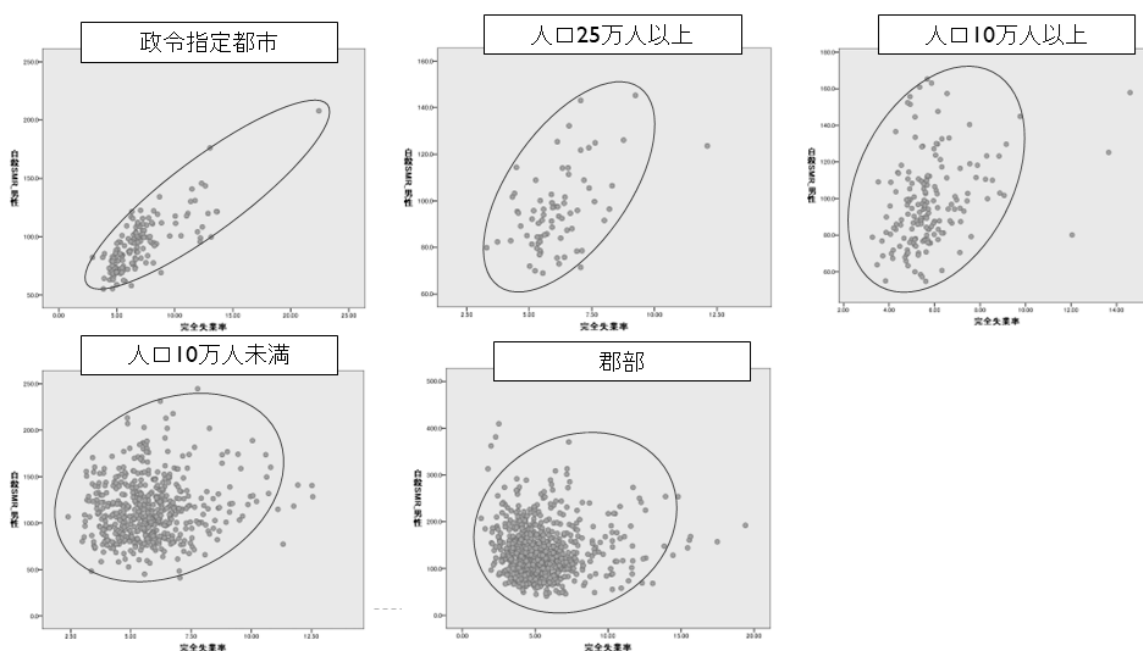


図 1 都市規模別にみた自殺 SMR(y)と完全失業率(x)の関係

いま、この図を前提に、新たな思考実験として、ある大都市で成功した自殺対策モデル

13 このことは、秋田モデルなどの既存の具体的モデルが他地域では成功しえないということを議論しているのではない。あくまで「ある特定の社会環境にあっては成功するモデルが、他の社会環境では必ずしも成功するとはいえない」という点に関する一般的議論である。誤解のないよう明記しておく。

14 この図は与謝野(2011)に所収の図の原図であり、第 84 回日本社会学会大会の報告資料で利用したものである。

15 実際には自殺に関する標準化死亡比(SMR)を利用している。SMR では、年齢の効果が簡便法で調整されており、全国平均を 100 として、そこからの乖離で自殺の多少を計測している。

を考えよう。その対策モデルは、大都市では失業率が自殺率と強い相関を持つことを前提に、失業による自殺発生の蓋然性を低下させるモデルとして大きな成功をおさめたものとする。このマイクロな対策モデルを、マクロな分析的な知識なしに郡部に適用したらどうなるだろうか。大きな効果を得る可能性が低いことは明らかだろう。なぜなら、郡部では失業率はキー変数ではなく、別の社会環境要因がより大きく影響している可能性が高いためである。このように、特定地域でのマイクロな成功モデルの他地域での適用は、その適用に先立ってのマクロな連関構造の知見がない場合、一般的には、実践的な有効性を持ちえない。

以上をまとめよう。前項で議論した通り、マクロな因果モデルを識別しても、自殺という個人的行為の直接的原因は原則的には解明できない。一方、自殺という個人的行為の直接的原因は、臨床医学的研究で地域特化的に解明されうるが、全国的な変化の理解には、マクロな要因分析が必須である。すなわち、マクロな要因分析を抜きにして全国的に有効な自殺対策は不可能であり、マクロな連関構造の識別は自殺対策の必須条件となる

自殺率の分析に関する以上の立場を基礎として、次節では、自殺率に対する新たな分析戦略の提案を行う。

2. 自殺率を規定する社会環境変数

2-1 自殺率に関する計量分析的アプローチ

自殺に関する既存の計量分析は極めて多く、ここで詳細なレビューを行うことはできないが、主なアプローチを以下の7つに整理した¹⁶。

1) 社会的統合論的アプローチ

デュルケームの伝統に則った議論であり、アノミー的自殺を中心に検討を進めるアプローチである。コホート分析等を適用するもの(Stockerd and O'Brien, 2002)、哲学・神学的、臨床的、心理学的、社会学的アプローチのそれぞれを対照しながら議論した実証分析(Jarosoz, 1998=2008)など計量的実証に関しても多数の論文がある。その他、離婚率、犯罪率など社会的統合の指標と考えられるものを分析対象とした論文は、アノミーの概念が明示されていない場合でも、このアプローチの中に位置付けてよいだろう。

2) 模倣論的アプローチ

デュルケームと対比的に議論されることの多いタルド(Tarde, 1890=2007)の模倣論的行動を扱うものである(Baller and Richardson, 2002)。英語圏で Copycat Suicide と表現される、有名人の自殺に刺激された自殺に関する研究もここに分類される(Niederkrötenenthaler et al. 2009; 野中 2015)。

16 ここでの6分類に収まらない論文もあり、また、ここで引用した論文も複数カテゴリーにまたがっているものがある。ここでは、主たる議論で暫定的に引用文献を6つに分類している。

3) 精神疾患アプローチ

臨床的アプローチの多くがこの中に入る。自殺未遂者の詳細な分析を行っているものも含まれる (Beautrais, Joyce and Mulder 1998)。医学系の計量分析ではこの立場から分析結果を考察しているものが多数ある (本橋・劉・佐々木 1999; 赤澤ほか 2010)。また、抑うつ論 (depression theory) アプローチと呼ばれる、飲酒・ドラッグ・若年層の性行動など、抑うつ精神状態を引き起こす疎外、逸脱行動に着目する議論もここに分類されるだろう (Morina and Duarte 2006)。

4) 経済的はく奪アプローチ

失業、GDP の変化、貧困率など経済的要因を中心に分析するものであり、失業による安定的な生活基盤の喪失を含め、経済的なはく奪が自殺を促進するとする (Freeman, 1998; 谷畑ほか 2003; Watanabe et al. 2006; Minoiu and Andrez 2008; 柴田 2014)。このアプローチの系として、生命保険の自殺への支払の免責期間が、自殺率をどのように規定するかといった分析も提出されている (Chen, Choi and Sawada 2008)。また、失業と健康問題という両者の喪失を同時に分析したものとして、岡ほか (2014) がある。

5) 社会移動論的アプローチ

階層的地位 (職業、所得) の世代間、世代内移動が、大きく自殺率に影響することを論じる。移動と統合の間関係が意識されている点では、社会的統合論の展開形ともいえる (Breed 1963)

6) コホート論アプローチ

コホートのサイズの大小が市場内競争圧力や政治的影響力に作用し、特定コホートの自殺率を左右するとするものである (Preston 1984; Easterlin 1987; Pampel 1996)。

7) 社会関係資本論的アプローチ

2000 年代以降に盛んな議論として、社会関係資本論的アプローチがある。このアプローチでは、人々の信頼感の醸成が自殺を抑止するとされ、コミュニティ内の信頼感の崩壊が自殺率を高めるとされる (Kawachi and Kennedy 2002; Kelly et al. 2009; 与謝野 2011)。この議論ではアノミー概念は中心的な位置を占めてはいないが、1) の社会的統合アプローチと類比的な構造をもっているといえる。

本研究では、上記の論文で利用されている変数群を参考に、分析データセットを構築する。

2-2 データ

2-2-1 分析単位と対象地域

本研究では、市区町村を単位としたデータを利用する。ここで、単位を都道府県としないのは、同じ県内でも大きな格差があり、それらをアグリゲートした都道府県単位では適

切な因果識別の可能性が低くなると考えられるためである。たとえば、兵庫県は南部に位置する神戸市を含む市部と中・北部に位置する郡部では、自殺率も社会環境変数も大きく異なる。都道府県単位の分析では、大都市部と郡部のデータがアグリゲートされ、内部の差異が平均化される。そのため、各都道府県内部に存在する自殺率と社会的環境変数の関係が無視され、結果、相関関係が希薄化されると想定できる¹⁷。また、Kunze and Anderson(2002)は、合衆国における分析で、州単位の固定効果を考慮すると既存の自殺の原因変数の多くが影響を失ってしまうことを指摘し、より小さな集団に分割して検討することが望ましいこと指摘している。これらの理由から、本研究では市町村を単位として分析を進める。

また、地域としては近畿圏に限定して分析する。この理由の第一は、パイロットスタディとして市区町村数を限定して効率的に分析を進めたいためである。第二の理由は、次節以降の分析結果をふまえ、個別の市区町村の状況を具体的にイメージした分析を容易にするためである。

2-2-2 分析データ

自殺に関するデータは、自殺予防総合対策センターの Web サイト「いきる」に掲載されていた「自殺死亡についての地域統計」を利用した。この地域統計には 5 年単位での自殺率のベイズ推定値を市区町村単位で掲載されている¹⁸。また、パイロットスタディとしての性格から、既存分析と対応するように、図 1 と同じ期間である 2003-2005 年を利用した¹⁹。

また、社会環境変数については、既存の自殺率の研究で利用されてきた変数のうち、政府が提供するデータポータルサイト e-stat において全国の市町村別データが提供されているものをピックアップし、まとめたものが表 1 である²⁰。いずれも 2005 年のデータを利用している。

次節では、自殺率とこれらの変数の組み合わせを試行錯誤し、分析上、有効と考えられる変数に絞って利用する。

17 ただし、都道府県単位でしか与えられていない地域内 GDP といったデータがあり、また、横断的・縦断的分析の混合モデルで市町村単位の分析が困難であることを考えると、都道府県単位での検討が有効あるいは必須である場合がある。

18 自殺総合対策センターが改組され、当該のホームページには現在アクセスが出来ない状態である。分析は過去にダウンロードしたデータを用いた。ちなみに、当時のアドレスは以下である。
(<http://ikiru.ncnp.go.jp/ikiru-hp/genjo/toukei/index.html>)

19 重複する年次の同様のデータが『自殺対策のための自殺死亡の地域統計 1973-2009』(自殺予防総合対策センター 2011)にも収録されているが、その後、いくつかの市町村でデータの修正が行われているため、こちらのデータは利用しない。

20 現状、すべてを網羅しているとは言えないが、かなりの程度、既存研究に対応するデータを整理した。また、当初、『民力 2015 CD-ROM 版』(朝日新聞社、2015)からデータセット作成を行っていたが、本研究で使用するデータの全てが e-stat に収録されていたため、e-stat に一本化している。

表 1 社会的環境変数リスト

人口動態	産業構成	インフラ	社会的統合	経済的安定・豊かさ
国勢調査人口	第一次産業就業者数	病院数	離婚件数	完全失業率
住民世帯数	第二次産業就業者数	高等学校数	婚姻件数	預貯金残高
一般世帯人員数	第三次産業就業者数	中学校数	刑法犯 認知件数	地方税収入額
15～64 歳人口率		小学校数		納税義務者 1 人当たり 課税対象所得
65 歳以上人口率		保育所数		税収入額伸び率
核家族世帯率				乗用車総保有台数
単身世帯比				
高齢夫婦のみ世帯率				
高齢単身世帯率				
転入人口				
転出人口				
転入超過率				

3. 近畿圏の全市区町村を対象としたパイロット分析

2-1 で挙げた先行研究は、いずれも理論的な視座からあらかじめ変数を選択し、そのうえで計量分析を行っている。仮説検証型の因果分析にけるオーソドックスな手続きといえよう。また、より探索的な分析の場合には、多数の変数をモデルに投入し、回帰分析における変数減少法や変数増加法と類比的な手法でモデルを識別していくというやり方もなされる²¹。ただし、両者ともに課題が残されており、前者の仮説検証型の分析手続きでは、仮説の視角外の変数の影響は当初より慮外とされてしまう。また、後者の手法は変数の選択が経路依存的で、適切なモデルの識別が困難である²²。

ここでは、以下の分析戦略に入る前に、予備的に自己組織化マップを利用し、変数をあらかじめ絞り込むことから始める。すなわち、表 1 の変数を自己組織化マップにかけ、出力される結果（U マトリックス）において、視覚的にクラスタが識別されていると判断されるまで変数を絞り込む。ここでは、絞り込みの結果を用いて分析戦略を進める。この絞り込みの妥当性をめぐっては、あらためて別稿で論じることとし、パイロットスタディとして、変数を絞り込んだ状態から以下で検討する。ここでは表 1 から 11 の変数に絞り込んだ。また、各変数は、適宜、人口、全産業就業者数、世帯数で除することで比率として扱っている。利用した変数は、65 歳以上人口率、核家族世帯率、高齢単身世帯率、第一次産業就業者率、第二次産業就業者率、第三次産業就業者率、人口 10 万人当

21 分散共分散分析において modification index でモデルを改善するといったやり方もあるが、検定統計量をもとに修正するという点では類比的な手続きといえよう。

22 交互作用まで考えると、多数の変数から適切な変数を選び出してモデル構築することは極めて困難と言える。

たり病院数、人口 1000 人当たり刑法犯認知件数、完全失業率、世帯当たり預貯金残高、
税収額伸び率の 11 変数である²³。

3-1 自殺率の分析戦略

本稿が提示する分析戦略を以下に示した(図 2)。また、次節で行う分析は、図 2 の Step6
までであり、Step7 以降は、個別の市町村の状況を詳細にみる必要があることから、別稿
で検討したい。



図 2 本稿での分析戦略の各ステップ

Step1 は、クラスタの識別を、自己組織化マップ(Self-Organizing Map: SOM)を利用し
て行うというものである。自己組織化マップの分析手法自体については、与謝野(2022)で
も簡潔に説明しているから割愛するが、多変数ベクトル間の距離を二次元上に投影する手
法であり、強い識別力を有する。ちなみに、ここで同様の性格をもつ多次元尺度法やクラ
スタ分析を用いないのには理由がある。

23 第一次～第三次の産業構成をそのまま分析に投入すると、完全な一次従属関係が生じてしまうた
め、自己組織化マップでは第一次産業就業者率を除いて分析し、平均の比較の際に再度計算しなおし
ている。

前者の多次元尺度法は、高次元の関係を低次元に縮約する古典的方法であり、その点で自己組織化マップと目標を共有している。ただし、多次元尺度法が全サンプルの相互類似度により低次元に落とし込んでいるのに対して、自己組織化マップは「参照ベクトルにより汎化して埋め込むため、大規模なデータに対する視覚的なデータマイニングに適した手法と考えられる」との指摘がなされている(福井ほか, 2005)。また、多次元尺度法が一般には三次元以上の空間でのモデルのフィットも想定する一方、自己組織化マップは二次元への投影に限定されており、視覚的な把握がより容易である。かつ、近接するように見えるクラスタでも、高い山によって切り離されるといった状況が視認でき、この点でも有利である。また、この近接性情報をもとに、クラスタを数値的に明確に切り分けることも可能であり、以降の分析戦略を構築しやすい。

後者のクラスタ分析については、多くの手法があるが、その中でも大きなケース数に対応する k-means 方が自己組織化マップと対比される。各ケース間の距離をもとめ、それによってクラスタに配分するという点で、両者の方法的近接性は高い。実際、古川(2017)は「SOM のアルゴリズムから協調過程をなくすと、k-means 法そのものになる。また学習量の分配に近傍関数を用いた Fuzzy k-means 法と見ることもできる」として、自己組織化マップと k-means 法の数学的対応を指摘している。ただし、基本的な k-means 方では、各ケースを 2 次元へ投影することが一般には困難であり、さらに、事前に指定された数だけ、特異点を含むクラスタを識別することになる²⁴。

これに対して、自己組織化マップでは、事前にクラスタ数を確定する必要がなく²⁵、さらには、特異点は他の点と距離のあるものとして描かれる。例えば、図の中心に特異点にあたるケース²⁶が位置していても、特異点は、高い山として描き出され²⁷、他のクラスタと隔絶していることが示される。このような描画をもとにすると、全ケースではなく、全体のうちの多数が含まれる複数のクラスタを識別することができる。

つまり、自己組織化マップを利用することで、全体のうちの大多数の傾向を数種類に判別でき、さらには、特異点に当たるケースを別途識別することができる。ただし、本稿では、特異点に当たるケースの識別までは行わず、大多数の傾向を把握するところまでを実践する。

また、自己組織化マップは、k-mean 法と同様に、各変数の単位が距離に反映してしまうため、分析に先立って、すべての変数の値を平均 0、分散 1 に標準化する。

さらに、ここでは自殺率を主要な変数としているため、自殺率が自己組織化マップの構

24 X-means 法と呼ばれる手法では、クラスタ数を統計的に推定し、自動的にクラスタ数を決定する。

ただし、事前の数学的仮定も強くなり、まだ、一般的でないことからここでは利用しない。

25 ただし、ノードと呼ばれるセル数の設定は必要になる。とはいえ、事前のクラスタ数の指定ほど制約は強くない。

26 実際にはケースそのものではなく、ケースに近接するノードと呼ばれるセル。

27 のちに示すように、高さは色の濃さで提示される。

成に与える影響(ノード間の距離に与える影響)を大きくするために、自殺率に 1.0 より大きなウェイト与える。ここでは、複数回の試行の結果、ウェイトを 2.0 とした。

Step2 では、与えられた周囲との距離を示す行列(U マトリックスと呼ばれる)を、近接性の基準を指定することで、「1or null」の行列に作り直す。一般には、自己組織化マップの分析結果を視認し、そこからクラスタを視覚的に識別するということが行われるが、ここでは、その手続きをより機械的に行う。このとき、周囲との距離がある基準よりも小さいノードのみを図示し、その区分をもとに次の Step の処理を行う。ちなみに、自己組織化マップでの識別がうまくいっているならば、適切な基準を試行錯誤することで、クラスタがきれいに切り分けられることになる。また、この基準の設定を、切り分けの明確さを示す指標を導入することで、自動化することも可能と考えているが、この自動化は今後の課題とし、ここでは基準値を試行錯誤で決定していく。

Step3 では、Step2 で得られた 1 の値が入ったノードについて、それぞれのプロフィールベクトル(各変数のノードにおける値)を作成し、それに対して k-means 法でクラスタを識別する。これは、Step2 で 1 を入れたノードを、目視で切り分けていく代わりに、k-means 法で自動的に切り分けようとするものである。その際、いくつのクラスタがあるかの設定は、エルボー法で行う。

エルボー法は、クラスタ数を 1 から順次増やし、SSE(Sum of Squared Errors of prediction)の減少傾向が大きく低下するクラスタ数をもって、全体のクラスタ数とするものである。Step3 はこのエルボー法の図示とクラスタ数の決定のステップである。

Step4 では、Step3 で設定されたクラスタ数で、k-means 法によるクラスタ分類を行い、Step2 で得られた 1 の入ったノードの布置を、クラスタ番号で塗り分ける。

Step5 では、塗り分けられたクラスタごとに、プロフィール(変数ベクトル)の平均を計算し、それぞれのクラスタの特徴を数値的に検討する。数値的に検討する際に、標準化されていた値をもとの単位に戻し、具体的な大きさとして認識できるようにする。

Step6 では、Step5 で対比的と思われるクラスタをとりあげ、その特性の違いをレーダーチャートで比較検討する。

Step7 では、各クラスタに含まれる市町村名のリストを作成し、それぞれの市町村の持つ特性を検討する。このことによって、平均で示されてきたものが具体的な名前をついたケースと対応し、クラスタの特性が具体的にイメージできるようになる。さらには、隠れている共通の値をとる変数(地理的特性など)や、差異を持つ可能性のある新たな変数の見極めが可能となる。

Step8 では、周囲との距離が大きく、いずれのクラスタにも含まれなかったケースを数値的、かつ具体的に検討する²⁸。

28 数値に加え、具体的な検討を同時に行うのは、特異値であるこれらのケースが少数であるとの想定からである。ただし、自己組織化マップの識別がうまくいっていない場合には、この限りではない。

本稿では、パイロット的に、上記の Step1～Step6 までを実践する。

3-2 分析戦略の実践

3-2-1 自己組織化マップ

前項 Step1 を行った結果が以下である²⁹。

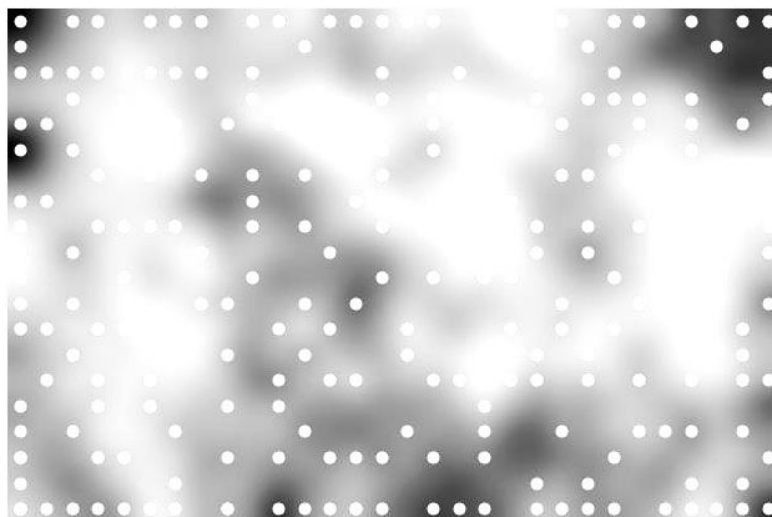


図 3 自己組織化マップの分析結果(Uマトリックスの図示)

ノード数を 20×30 で設定した。白抜きになって見える箇所は、ノード間の距離が小さいまとまりであり、クラスタにあたる。また、黒い部分は、他との距離が大きく、クラスタに隣接していても、クラスタとは特徴が大きく異なる部分である。また、“○”は実際に市区町村が存在するノードである。ただし、複数の市区町村が一つの“○”で表され、かつ、クラスタを構成する白抜き部分では“○”が視覚的に確認できないため、ここでは“○”の数、布置については問題としない。

図 3 から、黒い部分で切り分けられたクラスタがいくつか存在していることが見て取れる。通常は、この白い部分の視覚的な切り分けをここでおこなうが、先に記載した分析戦略にしたがって、Step2 へ進む。

3-2-2 Uマトリックスのクラスタ化

Step2 の手法で、各ノードと周りのノードとの距離を示したマトリックス(Uマトリックス)から、先に説明したように「1or null」行列をつくり、図 3 の白い部分を浮き出させるように処理する。白い部分は、周りのノードとの距離が小さい部分であるから、ある基準

29 分析には Python の somoclu ライブラリ ver.1-7-6 を利用した。また、Python 版 somoclu の設定で、ノード間がグラデーションされて描かれているが、各ノードは実際には“□”である。描画の設定を修正するには元のプログラムを修正する必要があるため、今回はプログラム設定そのまま実施している。

以下のノードを選び出せばよい。この作業はエクセルで自動化されている。

例えば、基準値を 1.2 とすると、図 4 のようなクラスタが浮かび上がる。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE		
1	1.2	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29		
2	0												1				1				1	1	1										
3	1																					1	1										
4	2					1																	1										
5	3				1		1					1	1	1						1	1				1								
6	4				1	1					1	1	1	1	1	1		1		1	1	1	1										
7	5				1	1	1	1						1	1	1				1	1	1	1				1	1					
8	6				1			1							1						1					1	1	1	1	1	1	1	
9	7				1										1			1	1	1	1	1				1	1	1	1	1	1		
10	8				1												1	1	1	1	1	1				1	1	1	1	1	1		
11	9	1			1			1	1											1	1	1					1	1	1	1	1		
12	10	1			1				1											1			1				1	1	1	1	1		
13	11				1	1																				1	1	1	1	1	1		
14	12			1	1	1	1	1	1				1				1						1		1			1	1	1	1		
15	13			1		1	1	1	1													1	1		1	1			1	1	1		
16	14			1		1																1			1	1			1	1	1		
17	15				1																	1								1	1		
18	16				1	1																									1	1	
19	17																															1	1
20	18							1																								1	1
21	19							1																								1	1

図 4 基準値を 1.2 としたときの「1 or null」行列

これが図 3 の自己組織化マップの分析結果の白い部分に対応している。一方、基準値を 2.0 とすると、図 5 のようになり、クラスタの分離が見えなくなる。

ちなみに、基準値の値は変数の数などによって異なるため、分析ごとに試行錯誤して決定することになる。基準値 1.2 はクラスタを浮き出させることには成功しているが、クラスタの範囲が小さすぎるきらいがある。基準値 2.0 ではクラスタの分離が確認できない。そこで、本分析では、各種の試行の結果、基準値 1.5 として分析を進めることとした(図 6)。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE
1	2	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
2	0			1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1			1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	5		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	6		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	7	1	1	1	1	1	1	1	1			1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	8	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

図 5 基準値を 2.0 としたときの「1 or null」行列

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE			
1	1.5	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29			
2	0					1		1		1	1	1	1			1	1	1	1	1	1	1	1	1										
3	1				1	1	1	1	1			1	1			1	1	1	1	1	1	1	1											
4	2				1	1	1			1			1	1			1	1		1	1	1	1	1										
5	3		1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1						
6	4			1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1					
7	5				1	1	1	1	1				1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1		
8	6			1	1	1	1	1					1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1		
9	7		1	1	1	1	1						1	1	1	1	1	1	1	1	1	1			1	1	1	1	1	1	1	1		
10	8	1	1	1	1	1	1	1	1					1	1	1	1	1	1	1	1	1			1	1	1	1	1	1	1	1		
11	9	1	1		1	1	1	1	1		1					1	1	1	1	1	1	1			1	1	1	1	1	1	1	1		
12	10	1	1	1	1	1		1	1	1						1	1	1	1	1	1	1				1	1	1	1	1	1	1		
13	11	1	1		1	1	1	1	1							1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
14	12			1	1	1	1	1	1	1		1				1	1	1	1	1	1	1			1	1	1	1	1	1	1	1		
15	13		1		1	1	1	1	1	1						1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
16	14			1	1	1	1	1	1	1					1			1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
17	15			1	1	1	1	1												1						1								
18	16				1	1	1																			1		1						
19	17					1	1	1																									1	
20	18						1	1	1																								1	
21	19	1						1	1																								1	1

図 6 基準値を 1.5 としたときの「1 or null」行列

3-2-3 クラスタ数の決定

Step3 では、エルボー法でクラスタを決定する³⁰。1~15 までクラスタ数を変えながら、SSE の変化を図示したものが図 7 である。図から、4、5、8 で折れ曲がりがあり、エルボーが出来ていることが見て取れる。通常では 4 あるいは 5 を取るべきとも思われるが、ここではより多くのクラスタ間の比較を行いたいこと、図 3 では 5 より多くのクラスタが見て取れることから 8 を採用した。

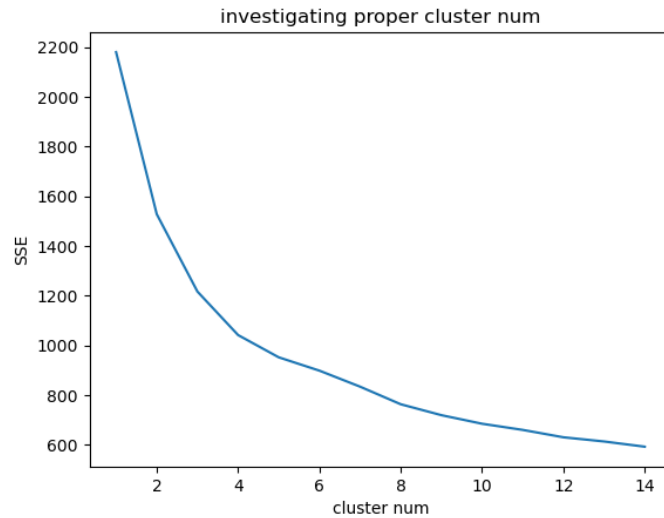


図 7 エルボー法によるクラスタ数の決定

3-2-4 k-means 法によるクラスタの塗り分け

Step4 では、エルボー法で決定したクラスタ数 8 と設定して、k-means 法によるクラスタ分析を行った。その結果を利用して図 6 を塗り分けた結果が図 8 である。

³⁰ クラスタ分析には Python sklearn ライブラリの KMeans クラスを利用した。また、SSE の変化も Python matplotlib ライブラリを用いて図示した。

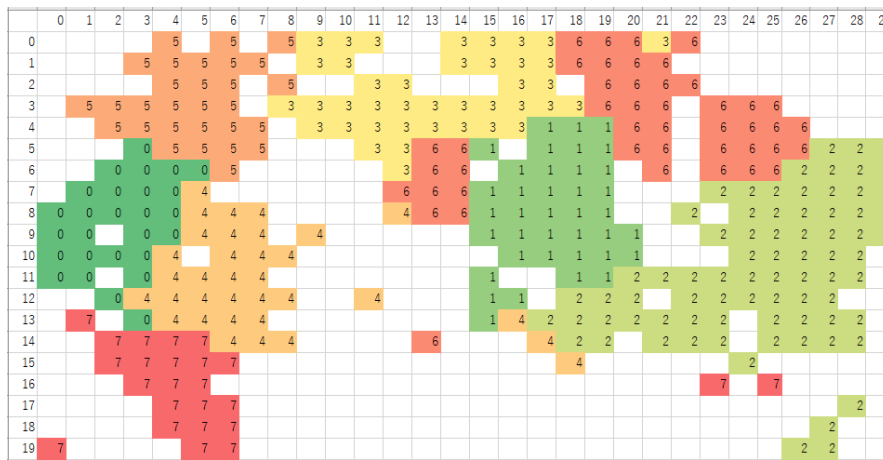


図 8 「1 or null」 行列のクラスタリング

図 8 を見ると、右下にクラスタ番号 2 の飛び地があるなど、各クラスタの飛び地が各所にみられる。自己組織化マップの原理からいうと、こうした飛び地があることは不自然ではあるが、クラスタが大きく分断されているといったことがないことから、暫定的に k-means 法の結果にしたがって、このクラスタ分割を利用して次のステップに進むこととする³¹。

3-2-5 各クラスタの平均

前項でもとめたクラスタについて、各変数の平均をもとの単位で示したものが表 2 である³²。表 2 をみると、クラスタ 7 が最も自殺率が高く、人口 10 万人あたりで約 47 人と極めて高率な自殺率を示している。このクラスタ 7 は第一次産業従業者が多く、高齢化も進んでいる地域であることが分かる。同じく第一次産業就業者が多いクラスタ 4 ではクラスタ 7 よりも高齢化が進んでいるが、自殺率は約 39 であり、全体の中では比較的高いとはいえ、クラスタ 7 よりもかなり低い値をとっている。この比較から、因果分析の独立効果だけを取り出し、「高齢化の進んだ農村部で、一般的に極めて自殺が高い」と結論することが危険であることが分かる。

また、クラスタ 2 は、犯罪率が高い地域であり、自殺率も約 40 と高くなっているが、同じく犯罪率の高いクラスタ 6 やクラスタ 1 では自殺率はそれほど高い値をとっていない

31 クラスタ 6 では、相対的に大きめの分離も起こっており、本来こうした飛び地を削除してから分析を行うことが望ましい。しかしながら、こうした結果がいくつか散見されるのは、自己組織化マップにおける使用変数が不十分、あるいは不適切であることを示唆しているから、より本質的な対応が必要とし、以後の課題としたい。このほかに、自殺率のウェイト、基準値、クラスタ数の変更の必要性を示唆しているとも考えられる。ここでは、パイロットスタディとしての性格から、これらの問題を理解したうえで、先のステップへ進むこととする。

32 Python では、連番の始まりが 0 となっており、ここでもそれを反映してクラスタ 0~クラスタ 7 までとなっている。

い。

このように、いくつかの変数の組み合わせなど、クラスタ間で自殺を規定する要因が異なる可能性が見て取れる。これらのクラスタ間の比較を行いながら、各変数の独立効果、交互作用の存在などを検討するのがここでの目的であり、表 2 はそうした検討が有意義である可能性を示唆する。ただし、こうした数値の対比ではクラスタ間の差の把握が困難であるから、次のステップでは、レーダーチャートを利用して視覚的に把握できるようにする。

表 2 各クラスタのプロフィール(オリジナル・スケール表示)

クラスタ	十万人当自殺数	第一次産業就業率	第二次産業就業率	第三次産業就業率	税収入額伸び率	六五歳以上人口率	核家族世帯率	高齢単身世帯率	完全失業率	世帯当預金残高 (百万円)	一人当病院数	十万人当犯罪数
0	41.01	5.57	38.25	56.18	103.31	21.35	56.85	6.47	4.81	6.04	6.03	15.84
1	34.98	4.77	29.90	65.32	101.42	17.56	67.61	5.99	6.29	4.57	4.17	18.65
2	39.82	3.60	29.38	67.02	101.47	19.08	64.37	8.66	8.17	7.03	6.25	22.26
3	28.30	4.33	26.27	69.40	101.48	17.48	69.45	5.94	6.07	6.79	4.19	14.26
4	38.90	15.19	28.86	55.95	101.47	29.14	55.68	10.75	4.67	4.54	8.00	10.43
5	32.74	6.40	36.35	57.25	103.50	18.78	58.89	5.38	4.51	7.95	4.08	17.12
6	32.83	4.06	22.98	72.96	102.41	18.73	64.44	8.07	6.29	8.93	5.60	19.18
7	46.94	13.82	26.99	59.19	101.76	25.78	57.61	9.29	5.35	5.46	4.67	11.02

3-2-6 レーダーチャートによる視覚的把握

クラスタ数は 8 であるから、一対比較の総数は 28 となる。そのすべてを検討することも可能であるが、ここでは自殺率が似ていながら異なる社会環境であるケース、ある社会環境が似ていながら自殺率が異なるケースをいくつか取り上げ、例示的な議論を行う。また、チャートの構成に際しては、単位の違いを無視できるように、表 2 を求める前の標準化した値で行う。また、図の変数名は略号で記載しているが、表 2 に対応して、12 時の位置から反時計回りに、自殺率、第一次産業就業率、第二次産業就業率、第三次産業就業率、税収入額伸び率、65 歳以上人口率、核家族世帯率、高齢単身世帯率、完全失業率、世帯当たり預金残高、人口 1 万人当たり病院数、犯罪率となっている。

まず産業構成が比較的似ている前述のクラスタ 4, 7 について図示する(図 9)。いずれのクラスタも、他の地域に比して、第一次産業構成比が極めて高い地域であることがわかる。ただし、赤で示されたクラスタ 7 は自殺率が極めて高い一方、青で示されたクラスタ 4 は比較的小さな値をとっている。また、高齢化率は青で示されたクラスタ 4 が高いが、高齢

単身世帯はクラスタ4の方が少ない。また、失業率はクラスタ4が低く、病院数についてはクラスタ4がかなり大きくなっている。

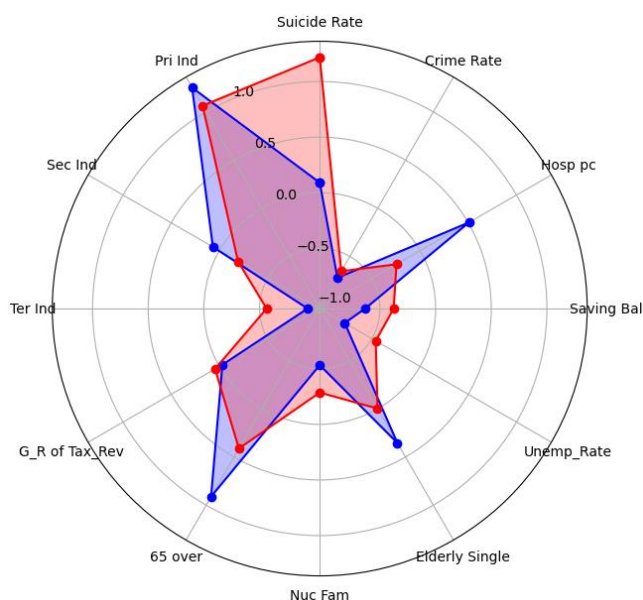


図9 クラスタ4(青)、クラスタ7(赤)のプロフィール比較

これまで「高齢者の自殺率が高い」ことは確認されている。また、社会環境変数の効果としては、既存研究で、「単身高齢者の方が孤独感による自殺が起きやすい³³⁾」ともされている。この2要因の効果を考えるとクラスタ4の方が、より自殺が多いという推測がなりたつ。一方、「失業率と自殺率は+の相関をもつ」という知見が比較的多いから、クラスタ4で失業率が低いことは自殺を抑止する方向に働く。この点で、高い高齢化率・単身高齢者率の自殺率上昇傾向を、低い失業率が抑えているという形になる。しかしながら、クラスタ4, 7のいずれも失業率は平均を大きく下回っており、失業率の差も大きくはない。このことから、クラスタ4で自殺率が低いことの主因が低失業率であるとは考え難い。

ところで、再度図を見直すと、クラスタ4が顕著にクラスタ7を上回っているのは、人口当たり病院数であることが分かる。このことから、人口当たり病院数の多さが、高い高齢化率・単身高齢者率が自殺率上昇させることを抑止し、結果、クラスタ7より低い自殺率に抑えている可能性が示唆される。この検討結果は結論的なものではなく、計量的因果モデルに先立つ基礎分析において、「人口当たり病院数を計量的因果モデルで利用することが有効性である可能性が示唆された」ということをのみを意味している。このことに注意したうえで、次の比較に進みたい。

クラスタ5, 6は、ほぼ同程度の低い自殺率を示している(図10)。

33 平光(2015)は、孤独感による自殺は独居群で高く、さらに男性、80歳以上で高いことを示している。

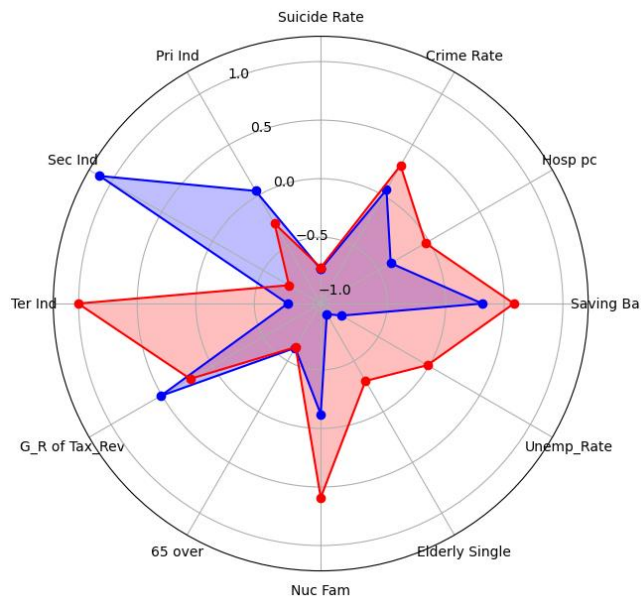


図 10 クラスタ 5(青)、クラスタ 6(赤)のプロフィール比較

核家族率、高齢単身世帯率、失業率、貯金残高、病院数、犯罪率のすべてにおいて高いのが、赤で示したクラスタ 6 であり、第三次産業従業率が極めて高いことから都市部と想定できる。一方、自治体の歳入の伸び率で若干上回る青線のクラスタ 5 は、第二次産業就業率が高いことから、工業を中心とする都市部と想定される。高齢化率の低さでは同一とはいえ、その他では、全く異なる二つの性格をもつ地域で自殺率がほぼ同一であることは興味深い。クラスタ 5 の際立って低い失業率、単身高齢世帯率がキーになっていることが示唆される。一方、クラスタ 6 の世帯の貯蓄残高の高さ、病院数が抑止効果を持っていることも示唆される。

次に、最も自殺率の低いクラスタ 3 と最も高いクラスタ 7 を比較する(図 11)。

ほぼ同一と言えるのが、第二次産業就業率、自治体の歳入の伸び率、病院数である。それ以外の変数では差がみられ、赤のクラスタ 3 は第三次産業中心であり、高齢化は進んでおらず、核家族率が高い。このことから、若年を中心とする世代によって構成される都市あるいはその周辺部と考えられる。失業率は若干クラスタ 3 が高いが、青のクラスタ 7 が高齢化の進む農村部であり、農業を行っている高齢者の失業が少ないと想定できることから、実質的な差は小さいと考えられる。また、世帯貯蓄残高も若干多いことから、若い世代としては比較的裕福な層が多い地域としてクラスタ 3 を位置づけられよう。すなわち、最も自殺率が高いのは高齢農村部であり、病院数も多いとはいえない過疎化が進みつつある地域と想定できる。一方、病院数は比較的少ないが、健康状態が総じて良い若い世代が経済的な豊かさのもとに生活している地域では、最も自殺が少ないと読むことができる。

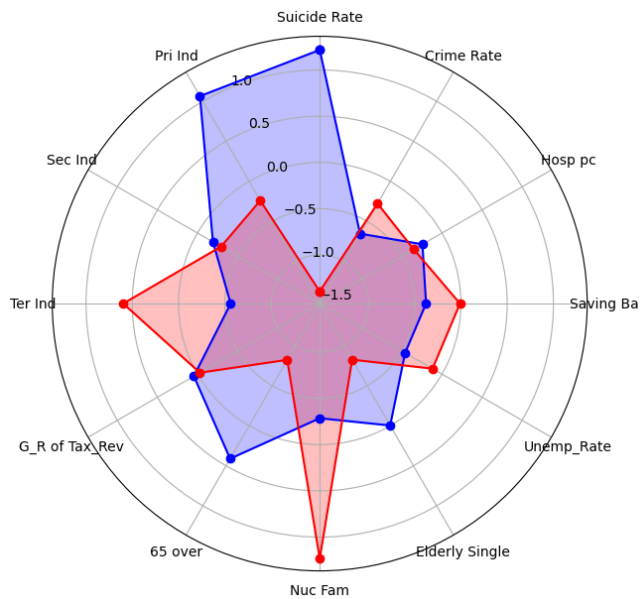


図 11 クラスタ 3(赤)、クラスタ 7(青)のプロフィール比較

また、クラスタ 3 の都市の若い世代という点で比較的類似しているのが、クラスタ 6 である。この両者について比較してみよう(図 12)。第 3 次産業就業率が高く、高齢人口が少ないなど全体の形は類似している。しかしながら自殺率には差がみられ、青で示したクラスタ 3 が低くなっている。失業率にも差がなく、貯蓄残高、病院数ではクラスタ 6 が勝っている点を考えると、この差は不思議にも思われる。一方、犯罪率、高齢単身世帯率でもクラスタ 3 が低いことから、こうした点が影響している可能性がある。すなわち、社会的統合の状態が比較的高く、高齢者の孤立が少ないことが、自殺率を抑止していると考えられる。つまり、社会的統合と孤立の様相の間に交互作用がある可能性が示唆されている。

最後に、クラスタ 7 に次いで自殺率が高い 3 つのクラスタ、すなわち 0、2、4 について比較を行いたい(図 13)。この 3 つのクラスタは、自殺率がそれぞれ約 41、40、39 であり比較的高い値を示している。図から明らかなように、緑で示したクラスタ 0、赤で示したクラスタ 2、青で示したクラスタ 4 はそれぞれ異なる形状を示しており、0 が工業地域、2 が都市部、3 が農村部に対応していると思われる。産業構成自体が自殺を促進するとは想定しにくいから、それぞれで相対的に突出した値をとっている変数に注目すると、都市部の 2 では「高-核家族率、高-失業率、高-犯罪率」、農村部の 4 では「高-高齢化率、高-高齢単身世帯率、低-貯蓄残高、高-病院数、低-犯罪率」で顕著な傾向を示している。他方で、0 の工業地帯では、自治体の歳入の伸び率が若干高いのみで、あとは突出した傾向を示していない。

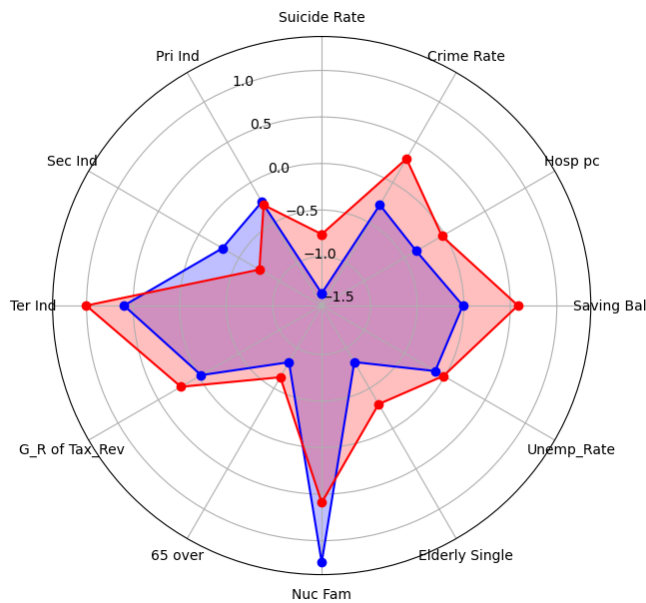


図 12 クラスタ 3(青)、クラスタ 6(赤)のプロフィール比較

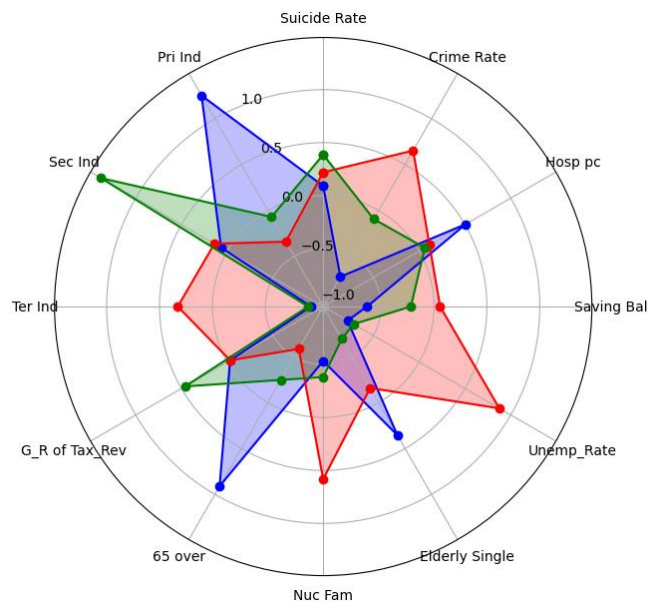


図 13 クラスタ 0(緑)、クラスタ 2(赤)、クラスタ 4(青)のプロフィール比較

いまここで、貯蓄残高を豊かさの指標とし、核家族化率・犯罪率を社会的統合の低さの指標とし、失業率を生活の不安定さの指標として解釈して状況を整理する。すると、都市部では、社会的統合が低く、生活基盤の安定性が失われたときに自殺が促進される地域が出現すると解釈できる。また、農村部では、低犯罪率からみて社会全体の統合は比較的高く、失業率から見た生活基盤の不安定さもないが、高齢化をめぐる諸問題と高齢世帯の孤

立化が自殺をめぐる主要な課題となっていると考えられる³⁴。都市部、農村部において、全く異なる社会環境要因が働いている可能性が示された。

一方、緑で示された工業地帯では取り分けて問題になるような値がなく、突出した特徴も示されていない。自治体の歳入の伸び率において高い値を示してはいるが、これが自殺の促進要因であるとは考え難いことから、ここに示されない変数の効果が大きいこと、すなわち、新たな変数の投入必要性が露わになってきている。また、世帯の貯蓄残高が平均より低く、失業率および核家族率が農村部とほぼ同一であることの交互作用がこうした平均より高い自殺率を生み出している可能性もある。隠れた説明変数や交互作用の存在可能性を検出することは、今回の分析戦略の大きな目的の一つであるから、この点で、クラスタ0のプロフィールも極めて示唆的である。

4. 全体のまとめと今後の課題

本稿では、自殺率に関する計量的因果モデル分析に先立って、適切な変数群の識別を行うための分析戦略を提示した。ここで提示した分析戦略は、自己組織化マップと k-means 法を段階的に用いることで地域のクラスタを識別し、各クラスタのプロフィールを検討することで各変数の効果を識別しようとするものであった。この分析戦略を用いた本パイロットスタディの結果は、第一次産業、第二次産業、第三次産業のそれぞれが主要となって地域で、自殺率を規定する要因が異なることを示唆していた。この結果は与謝野(2011)の分析結果と対応するものであり、地域特性、社会環境によって自殺対策の有効性が異なる可能性が高いという本稿の理論的整理とも対応するものであった。

第一次産業が中心である都市部では、社会的統合が低く、失業によって生活基盤の安定が失われたとき自殺が促進されると想定できる。また、第三次産業が中心である農村部では、社会的統合の程度は相対的に高いが、高齢化に伴う諸問題(心身の健康問題など)が高い自殺率を生み出すと想定される。ただし、1-2で述べた通り、個人的行為としての自殺をめぐる状況を明らかにすることは、マクロな自殺率の分析からは不可能であるため、今後の自殺対策の指針を与えるに過ぎないことは、ここであらためて明記しておきたい。

また、第二次産業が中心であるクラスタ0では、自殺率を上昇させるような突出した単独要因が見いだせず、交互作用要因や未知の変数の効果などが想定できる。このクラスタでは、クラスタ2の都市部と似た社会環境とクラスタ4の農村部と酷似した社会環境の両者を有している。世帯の貯蓄残高が平均以下であり、農村部と共通するような低い失業率と核家族率の組み合わせが比較的高い自殺率を生み出している可能性が示唆される。また、今回の分析では投入されていない変数の効果が隠れている可能性もある。そのた

34 ただし、単身高齢世帯率の高さは、高齢化に比例して必然的にあがるため、単身世帯でない世帯の自殺がより大きく、単身高齢世帯率の高さの独立効果は逆向きである可能性もある。この点については、単身高齢世帯率の変数の再構成(高齢化率との比の計算など)の必要がある。

め、変数群を入れ替えながら、ここで提示した分析戦略の Step1～Step6 までを根気強く繰り返す、クラスタの特徴を描きなおしながらより明快な解釈へと近づけていく必要があるだろう。

本分析では Step7 の地域名とクラスタを対応させ、地理的情報を含む多様な情報からクラスタの具体的なイメージを構成するといった、いわば質的な分析のパートを割愛した。また、Step8 に対応する図 3 で濃い黒で塗られた地域についても分析を割愛している。図 3 の濃い黒の部分は、多数の地域を含むクラスタを構成することはなくとも、特徴的な地域として重要であり別途の分析戦略が必要である。これらの点に関しては、改めて分析・検討を行うこととしたい。ところで、今回の自己組織化マップをめぐる分析戦略にはいくつかの課題がある³⁵。

第一に、自殺率に対するウェイトの問題がある。本分析の目標は、自殺率に対する社会環境変数の影響の予備的整理にある。そのため、自殺率以外の変数の近さ、遠さに大きく依存して識別されたクラスタでは、クラスタ内部で自殺率の分散が大きかったり、自殺率では類似した地域が別クラスタとなったりしてしまうといったことが起こる。もちろん、クラスタの特徴を記述するとき、各変数の平均を利用して解釈をしているから、社会環境変数に関しても、クラスタ内の分散が大きすぎる場合にはクラスタの解釈に問題が生じる。自殺率のウェイトを大きくすると、自殺率によって各ノードの距離がより大きく影響されやすくなり、小さくすればその逆に自殺率の影響力は小さくなる。この適切なウェイトをどのように決めるのかという問題がある。変数の数によってもウェイトの効果は左右されるため、一概に基準を決めることが難しい。現行では、図 3 のような U マトリックスでクラスタが明確に分かれているかどうかを判定し、試行錯誤で行っている。この点の形式的手続きの構築は今後の方法論的課題である。

第二に、図 8 のように、飛び地がある図としてクラスタが描き出されたときの対処の仕方である。自己組織化マップでは、ノード間の距離が遠いほど、その距離は加速度的に増加するように設計されている³⁶。かつ、図 3 で黒く描かれた部分は、他との間を隔てるような周りとの距離が遠い部分であり、地図で例えるならば急峻な山にあたる。すなわち、白く描かれた複数の平野部を山が隔てているかのように図 U は描かれている。そのため、自己組織化マップでは飛び地のように、距離の遠い部分で同一のクラスタが構成されることは基本的にはあり得ない。今回はそのことを理解したうえで、その影響が小さいと想定して分析ステップを先に進めた。しかしながら、本来的には、こうした飛び地を除いて分析すべきであり、さらに言えば、変数の調整などの再分析によって U マトリックスを再構成し、飛び地がないような分析結果を得る必要がある。ちなみに、飛び地を自動

35 ノード数の決定に関する問題もあるが、これは自己組織化マップ一般の問題あるからここでは特に取り上げない。

36 一般に指数的に距離が増加するように設計される。

的に除くことはプログラム上可能であり、今後、そうしたシステムの構築についても試みることとしたい。ただし、より重要なことは、飛び地がないような分析結果を得ることと考えており、今後、変数の調整など各種の改善を試みる。

最後に、今回の分析では平均にのみ着目し、クラスタ内部での各プロフィール変数の分散を考慮していないという問題がある。すなわち、クラスタ間の比較において級内分散、級間分散を考察するといったことがなされていない。このことは、クラスタ間の差についての効果量を測定していないということでもある。全数である今回のデータで統計的検定を行うことに関しては各種の議論があろうが、効果量を測定することは比較の上で重要である。また、統計的検定に関しても、慎重かつ適切な検討を行ったうえで、必要があれば実施すべきであろう。これらの点に関しては方法的な議論を整理したうえで、今後の重要な課題としたい。

上述のように、本パイロットスタディには複数の残された課題がある。しかしながら、今後、自殺率に関する高い予測力を持つ因果モデルを構成するために、本分析戦略が一定の意義をもつことは示されたものと想定している。今後、本分析戦略を用いて、時系列での比較を行うなどの展開を行い、さらなる手法的彫琢を加え、いずれ高い予測力を持つ因果モデルの構成を目指すこととしたい。

引用文献

- 赤澤正人・松本俊彦・勝又陽太郎・木谷雅彦・廣川聖子・高橋祥友・川上憲人・渡邊直樹・平山正実・亀山晶子・横山由香里・竹島正, 2010, 「死亡時の就労状況からみた自殺既遂者の心理社会的類型について」『日本公衆衛生雑誌』 57(7): 550-560.
- 秋田県自殺対策計画策定委員会, 2023, 『第2期秋田県自殺対策計画(素案)』 秋田県庁
- Baller, R. D. and K. K. Richardson. 2002, "Social Integration, Imitation, and the Geographic Patterning of Suicide," *American Sociological Review*, 67: 873-888.
- Beautrais, A., P.R. Joyce and R. T. Mulder, 1998, "Unemployment and Serious Suicide Attempts," *Psychological Medicine*, 28: 209-218.
- Breed, W., 1963, "Occupational Mobility and Suicide among White Males," *American Sociological Review*, 28(2): 179-188.
- Chen, J., Yun J. Choi and Y. Sawada, 2008, "Suicide and Life Insurance," *CIRJE Discussion Papers*, 558:1-21.
- Durkheim E., 1897, *Le suicide: étude de sociologie*, Paris, Félix Alcan (J. A. Spaulding and G. Simpson trans., [1952] 2002, *Suicide: A study in sociology*, Routledge Classics, Routledge).
- Easterlin, R. A., 1987, *Birth and Fortune: The Impact of Numbers on Personal Welfare*, Chicago: University of Chicago Press.

- Freeman, D., 1998, "Determinant of Youth Suicide: The Easterlin-Holinger Cohort Hypothesis Re-examined," *American Journal of Economics and Sociology*, 57(2): 183-199.
- 藤田利治, 2009, 『自殺死亡についての地域統計』国立精神神経センター精神保健研究所・自殺予防総合対策センター.
- 福井健一・斉藤 和巳・木村 昌弘・沼尾 正行, 2005, 「自己組織化マップによる教師情報を用いた可視化アーキテクチャの提案--時系列医療データの可視化を例に」『情報処理学会研究報告』78: 17-22.
- 古川徹生, 2017, 「自己組織化マップ入門」, Furukawa Laboratory ホームページ, (2023年3月1日取得, <http://www.brain.kyutech.ac.jp/~furukawa/data/SOMtext.pdf>).
- 平光充良, 2015, 「孤独感による自殺志望と同居人の有無の関連」『厚生学の指標』62(6): 16-19.
- 池田一夫・伊藤弘一, 1999, 「日本における自殺の精密分析」『東京都衛生研究所年報』50: 337-344.
- Jarosz, M., 1998, *Suicide*, Institute of Political Studies Polish Academy of Sciences. (石川晃弘・石垣尚志・小熊信訳, 2008, 『自殺の社会学—ポーランド社会の変動と病理』学文舎.)
- 自殺予防総合対策センター, 2011, 『自殺対策のための自殺死亡の地域統計 1973-2009』国立精神・神経医療研究センター精神保健研究所.
- 河西千秋, 2008, 「救命救急センターにおける自殺未遂者への支援と自殺再企図予防方略の開発」『学術の動向』13(3): 39-43.
- 河西千秋, 2011, 「自殺問題が深刻化してから以降のわが国の自殺未遂者対策の経緯」『総合病院精神医学』23(3): 241-246.
- Kawachi, I. and B. Kennedy, 2002, *The Health of Nations: Why inequality is harmful to your health*, New York: The New Press. (西信雄・高尾総司・中山健夫 監訳, 2004, 『不平等が健康を損なう』日本評論社.)
- Kelly B. D., M. Davoren, Á. N. Mhaoláin, E. G. Breen and P. Casey, 2009, "Social Capital And Suicide in 11 European Countries: An Ecological Analysis", *Social Psychiatry and Psychiatric Epidemiology*, 44: 971-977.
- 紺田広明・与謝野有紀, 2013, 「コホートサイズの自殺率規定力に対する年齢効果の検討—世代間, 世代内コンフリクトが日本の自殺率に対して与える影響の計量分析—」『社会的信頼学』1: 23-42.
- Kunze, M. and A. L. Anderson, 2002, "The Impact of Socioeconomic Factors on State Suicide Rates: A Methodological Note," *Urban Studies*, 39(1): 155-162.
- Martikainen, P. T. and Volkonen, T, 1996, "Excess Mortality of Unemployed Men and

- Women during a Period of Rapidly Increasing Unemployment,” *Lancet*, 348: 909-912.
- 松本寿昭・若林佳史・小森田龍生・小牧奈津子・松山博光・安田和子・田所満理奈・反町吉秀, 2013, 「予防に向けた自殺の要因に関する研究-世代・文化・コミュニティの視点から-」『人間生活文化研究』 23: 198-202.
- Minoiu, C. and A. R. Andrez, 2008, “The Effect of Public Spending on Suicide: Evidence from US State Data,” *The Journal of Socio-Economics*, 37: 237-261.
- Morina, J.-A. and R. Duarte, 2006, “Risk Determinants of Suicide Attempts among Adolescents,” *Journal of Economics and Sociology*, 65(2): 407-434.
- 本橋豊・劉揚・佐々木久長, 1999, 「『秋田』県の自殺志望の地域較差と社会生活要因に関する研究」『厚生指標』 46(15): 10-15.
- Niederkrötenhaller, T., B. Till, N. D. Kapusta, M. Voracek, K. Dervic and G. Sonneck, 2009, “Copycat Effects after Media Reports on Suicide: A Population-Based Ecologic Study,” *Social Science and Medicine*, 69(7): 1085-90.
- 野中昭彦, 2015, 「著名人の自殺後に見られる後追い自殺 -時代の推移による変化-」『流通科学研究』 14 (2): 105-115.
- 岡檀・久保田貴文・椿広計・山内慶太, 2014, 「日本の自殺上昇期における地域格差に関する考察」『厚生指標』 61(8): 8-13.
- Pampel, F. C., 1996, “Cohort Size and Age-Specific Suicide Rates: A Contingent Relationship,” *Demography*, 33(3): 341-355.
- Preston, S. H., 1984, “Children and the Elderly: Divergent Paths for America’s Dependents,” *Demography*, 21: 435-457.
- 佐々木洋成, 2005, 「アノミーの社会史－性、年齢、地域別自殺死亡率の経年分析」『社会学評論』 55(4): 468-482.
- 佐々木隆一郎, 2006, 「マクロとミクロ」『信州医誌』 54: 115-116.
- 谷畑健生・藤田利治・尾崎米厚・黒沢洋一・蓑輪眞澄・畑栄一, 2003, 「自殺と社会背景としての失業」『厚生指標』 50(8): 23-29.
- 柴田悠, 2014, 「自殺率に対する積極的労働市場政策の効果-OECD26 ャ国 1980～2007 年のパネルデータ分析」『社会学評論』 65(1): 116-133.
- Stockerd, J. and R. M. O'Brien, 2002, “Cohort Effects on Suicide Rates: International Variation,” *American Sociological Review*, 67: 854-872.
- Tarde, J.-G., *Les lois de l'imitation: etude sociologique*, Paris: Félix Alcan. (池田祥英・村澤真保呂訳, 2007, 『模倣の法則』河出書房新社.)
- Watanabe, R., M. Furukawa, R. Nakamura and Y. Ogura, 2006, “Analysis of the Socioeconomic Difficulties Affecting the Suicide Rate in Japan,” *KIER Discussion Paper*, 626: 1-19.

与謝野有紀, 2011, 「格差、信頼とライフチャンスー日本の自殺率をめぐってー」 斎藤友里子・三隅一人編著『現代の階層社会 3 流動化のなかの社会意識』 東京大学出版, 293-307.

与謝野有紀, 2022, 「階層問題としての団地高齢化ー自己組織化マップを用いた予備的検討」『関西大学社会学部紀要』 53(2): 1-22.

Proposal of a Novel Strategy for Identifying Independent Variables for Causal Analysis of Suicide Rates : Pilot Study Applying Self-Organizing Maps

Arinori Yosano

[Abstract]

Suicide is one of the social pathological phenomena that is exacerbated by a lack of social trust. In Japan, there was a sharp increase in the number of suicide cases in 1998, and over 450,000 people died by suicide between 1998 and the following 14 years. Although the number of suicides has remained in the 20,000s since then, there has been a recent increase in male suicides in 2022, the first rise in 13 years, and the number of suicides among elementary and secondary school students has exceeded 500 for the first time. Thus, suicide remains a pressing issue in the present day. Numerous quantitative analyses have been conducted to explore the reasons behind the fluctuations in Japan's suicide rate. However, there is no consensus among researchers regarding the cause of the sudden surge in the suicide rate in 1998. Additionally, there has been no definitive causal model for the fluctuations in suicide rates. To address this research situation, this study proposes a novel analytical strategy. Specifically, prior to conducting causal analysis, the study utilizes self-organizing maps to cluster regions and efficiently identify groups of variables. The analysis consists of eight steps, and in this study, the first six steps were carried out using actual data, demonstrating the effectiveness of this new analytical strategy.

Keywords and Phrases: Suicide, Suicide Rate, Self-Organizing Map, Strategy for Identifying Independent Variables