

計量社会学と因果推論：

観察データに基づいた社会の理解に向けて

筒井 淳也（立命館大学）

【要約】

統計的因果推論は計量分析の主流となっているが、計量社会学におけるその意味やインパクトについて体系的に論じた研究はいまだに少ない。本論文では、因果推論モデルを含む計量分析の手法について、異質性という概念を軸に整理し、その上で計量社会学が異質性に対して他の分野の手法とは異なったアプローチをとる傾向があることを示す。このことは、マルチレベル分析とも呼ばれる混合効果モデルの活用において明らかである。さらに、介入や切断を用いる因果推論アプローチと、要因間の関連性を強調する計量社会的アプローチの違いを説明し、それが人々の概念連関を参照する社会学の特性の現れである、ということ論じる。

【キーワード】

因果推論, 計量社会学, 観察データ

1 はじめに

因果推論は、すでに計量分析の柱のひとつになっている。ただ、「統計的因果推論」という場合、大きく分けて二つの潮流があるようだ。ひとつは、多くの場合「潜在的アウトカム (potential outcome)」アプローチと呼ばれる流れである (Rubin 1974, 2005)。この場合、たいていは一つの原因を「介入 (intervention)」として想定し、その結果の大きさを効果として考える。もう一つは、複数の原因と結果の関係を「因果ダイアグラム」としてモデル化し、構造方程式等で推定を行うというアプローチである (Pearl 2000=2009)。

前者の因果効果の推定は、理想的には介入実験データによって行われるものだと考えられているが、調査観察データにおいて因果効果の推定を試みるための代表的な手法として、マッチング分析や傾向スコア分析を挙げることができる。一連の因果推論モデルは、手法を体系的に示したドナルド・ルービンの名前をとって、しばしばルービン・モデルと呼ばれている。本論文でも、調査観察データに対して因果推論を試みるこれらの手法をルービン・モデルと呼んでおく。後者のアプローチは、ルービンではなくむしろジュディ・パールの研究を参照し、伝統的計量社会学でも馴染み深いパス解析と親和性が高い。

もちろん、これらの二つの立場が根本的に異なった因果の捉え方をしていたり、あるいは因果推論に対する著しく異なったアプローチをしているわけではない。ただ、両者は異なった概念群を用いることが多いし、一方に依拠したテキストで他方の重要研究があまり参照されない、といった傾向性もある。

たとえば介入概念については、ルービン・モデルにおいては無作為化比較実験の語彙で論じられるが、パールの提唱する因果ダイアグラム（因果ベイズネット）においては、「グラフ切除（graph surgery）」（Pearl 2000=2009）という言葉が使われているように、グラフ理論に関連した捉え方をされる。教育用テキストについても、アメリカの社会学教育において参照されることが多いと言われている Morgan & Winship（2007）ではルービンの文献は10個以上引用されているが、パールの文献は一つだけである¹⁾。逆に、国内でのグラフ理論を活用した因果推論のテキストである宮川雅巳（2004）においては、パールの研究の引用が多く、ルービンはひとつだけである。また、両派のあいだには論争もあった。発端は、パールによる「ルービンの potential outcomes アプローチには弱点がある」という指摘である（Pearl 2000=2009: 350）。その後両派のあいだで、*Statistics in Medicine* 誌上でのやりとりが続いた。

本論文では、これらの関係自体について深く考察することはしない²⁾。ただ、以下では、因果推論アプローチとして主にルービン・モデルを想定し、それが伝統的手法との関係において持つ含意について検討する。

視野を広くとってみれば、マッチングに代表されるルービン・モデルは、因果推論を「より強く」志向する一連の分析手法だということができる。「より強く」というのは、それ以前のオーソドックスな計量分析手法、特に回帰分析においても、分野によっては因果性が明確に意識されてきたからである。回帰分析の推定量が有効かつバイアスのない結果を返す条件（OLS 推定量が線形最良推定量 Best Linear Unbiased Estimation: BLUE である条件）については、計量経済学の分野で体系化されて久しい。2000年代においてはまだ、計量経済学の入門テキストにおいてルービン・モデルは「応用的手法」のひとつとして、しばしば注釈的な位置づけになっていることもあった³⁾。むしろ注目を集めていたのは、自然実験データ、観察データにおける操作変数法や回帰分断モデルの活用であったといえよう⁴⁾。

ただ、いずれにしろ計量経済学の分野では、計量社会学におけるよりもバイアスのない因果推論を追究する傾向が強かったことは否定できないだろう。計量経済学や、実証分野においても労働経済学、開発経済学で自然実験データを重用する研究グループが目立つということを考えても、「伝統的」な計量経済学とルービン・モデルの相性は極めて良いということができる。また、ルービン自身が心理学分野の出身であることにみてもとれるように、心理学においても因果推論のモデルはそれほど違和感なく溶け込んでいるとみるべきであろう。

他方で、計量社会学、あるいは社会学における数量的実証研究においては、強い意味での因果推論志向のモデルがそれほど重視されなかった。古典的回帰分析はすっかり根付いているものの、計量経済学分野におけるように自然実験データを探索するケースは目立ったものではない⁵⁾。

この現状から引き出しうるインプリケーションにはどういったものがあるだろうか。以下、本論文では、計量社会学と因果推論との関係を、異質性（heterogeneity）という概念を軸として整理したうえで、「人々の概念連関への参照」という観点から計量社会的なアプローチの特性について議論する。その際、主に計量社会学分野で隆盛をみたマルチレベル分析（混合効果モデル）を因果推論アプローチと対比し、計量社会学の特性について考察するきっかけとする。それを受けて、計量社会学がデモグラフィック変数による「記述」に一定の価値を与えていることを明らかにする。最後に、因果推論モデルが計量社会学にとって持つ含意について述べる。

本論に入る前に付言しておく、この論文は、分野ごとに異なったアプローチがある

「べきだ」と主張するものではない。そうではなく、分野ごとに異なった強調点が事実問題として観察され、かつそういった相違は分野の研究特性に関連しているのではないかと考えるものである。

2 各種計量手法における個体の異質性の位置づけ

2.1 デモグラフィックデータと個体の異質性

計量社会学においてルービン・モデルに代表される因果推論の手法の導入が遅れている、あるいは自然実験データの活用にあられるような、「強い」因果推論志向がそもそも欠如しているとすれば、そういった現状からどういったインプリケーションを汲み取ることができるだろうか。

ひとつには、単純に「心理学や経済学などの他分野に比べると遅れているから、もっと普及させよう」という方向性であろう。この主張は当然、有効である。というのも、すでに述べたように計量社会学の実証研究において以前から因果推論志向が存在していることも間違いではないからだ。ただ、たとえばOLS推定量が本来満たすべき条件について、社会学では経済学ほど厳しい配慮がされることはなかった。確かに、ルービン・モデルが対応できる問題は、異質性がもたらすバイアスに対してはほんの一部に限られる。ルービン・モデルの主要な機能は、説明変数ごとのデータの異質性を、観察されたその他の説明変数（統制変数）で可能な限りバランスさせる、という手続きにある。しかし、非常にシンプルなこの手続きさえ、計量社会学において一般に共有されているとは言い難い。本来はより「強い」因果推論の手法を使うべきところで、そうではない手続きが使われているのである。

ただ、ここではもうひとつのインプリケーションの方がより重要であると思われる。それは、計量社会学がいわゆるデモグラフィック要因（個体の異質性）に強い関心を払うがゆえに、個体の異質性の効果を除去することを目的とするルービン・モデルやその他の因果推論志向が強い手法を忌避したか、あるいは導入の必要性が強く意識されなかった、という可能性である（Xie 2007）。

デモグラフィック要因あるいはデモグラフィックデータとは、人々のライフコースにおいて全く変化しないか、あるいは変化しにくいものを指している。出生年代、出身地、特定時点での親の職業といった変数は、本人の人生において、他の要因の影響を受けて変化することがまったくない。こういったデータが「変わる」とすれば、計測エラーや、解釈上のカテゴリー変更があった場合に限られる。また、性別、学歴、職業といったデータも、比較的变化しにくいものであろう。比較的固定的であるがゆえに、これらの変数は同一個体（しばしば個人）内において安定している固有の性質である。このような意味合いを込めて、デモグラフィック（人口学的）なデータだと考えられているのである。そして計量社会学においては、いわゆる「地位達成モデル」にみてとれるように、デモグラフィックデータを他のデモグラフィックデータによって「説明」という研究がしばしばなされてきた。

一方、マッチングに代表されるルービン・モデルにおいては、基本的にはデモグラフィックな異質性の影響をいかにして消去するかが課題となる。というのは、個体の異質性こそが、（しばしば）セルフ・セレクション（＝自己選択）による説明変数の割当を通じて、結果に歪みをもたらすと考えられているからである。個体の異質性の多くは観察できない

傾向性を生み出すと考えられているため、説明変数と被説明変数の両者に影響を与える可能性がある。もし、自己選択が観察されやすいデモグラフィック要因によってのみもたらされるのならば、それに起因するバイアスに対しては、ルービン・モデルで対応可能である。そうではない場合、自然実験や操作変数法といった手続きが要請される。

このように、個体の異質性をどのように扱うのかについて、計量分析は一定の広がりを持っている。以下では、調査観察データの分析を前提としつつ、計量分析が異質性をどのように位置づけているのかという観点から、計量社会学で普及している手法を含めて、いくつかの計量手法を整理してみよう。

2.2 計量経済学における異質性の位置づけ

ジェイムズ・ヘックマンがノーベル経済学賞受賞記念講義で述べているように、計量経済学の歴史は、個体の異質性ととの格闘の歴史でもあった。「回帰分析についていえば、切片のみならず傾き係数が、あるいはその両方が説明変数と関連しうる。異質性と多様性、それらが経済学と計量経済学にとって持つ含意は、この講義と私のライフワークにとっての中心の課題なのだ」(Heckman 2001: 674-75)。

計量経済学において伝統的に好まれる推定量の性質は、いうまでもなく不偏性 (unbiasedness) (あるいは一貫性 (consistency)) と、有効性 (efficiency) あるいは誤差の適正な推定である。個体の異質性は、この両者に深く関係する。

古典的最小二乗推定量では、メインの説明変数以外の説明変数 (主要なデモグラフィック変数) を統制することで、不偏推定値を目指す。 $y = X\beta + \varepsilon$ という回帰式において、 $\text{Cov}(X, \varepsilon) = 0$ 、すなわち説明変数と被説明変数との両者に相関する要因が ε (誤差項) のなかに含まれていないことが前提となる。さらに、誤差が適正に推定されている条件は、 ε の分布が i.i.d (分散均一で互いに独立) である、ということだ。

観察された説明変数の効果は、回帰分析の語彙において「固定 (fixed)」(繰り返しサンプルにおいて固定的) であるという前提があるために、しばしば「固定効果 (fixed effects)」と呼ばれる。これに対して、誤差項に含まれる個体の観察されない異質性はしばしば「変量効果 (random effects)」と呼ばれる。古典的最小二乗推定量の基本的な分析方針は、関心のある要因とともにデモグラフィックデータを調査観察において同時に採取し、モデルに投入することによって、変量効果に含まれる、バイアスをもたらし未観察要因を観察された固定効果に移し入れる、というものである。同時に、誤差相関に起因する誤差の過小推定については、ロバスト推定量や GLS、変量効果最尤推定といった「メニュー」が準備されている。

個体の異質性の影響を、個体のデモグラフィックなデータの観察を通じて除去しようという方針は、個体のセルフ・セレクションをモデリングしたヘックマン推定量や、マッチング推定量、傾向スコア推定量などのルービン・モデルにおいても共通している。固定効果推定量においては、同一個体の反復観察を通じて、個体変数を投入、あるいは個体間の変動 (between データ) をキャンセルすることで、個体内の変動 (within データ) のみを用いて回帰を行う。バイアスをもたらし要因を理論的に推察し、それと (理論的に) 相関しない変数を観察・投入することで不偏推定を試みる操作変数法 (あるいは心理学分野における媒介変数法) も、観察データによる調整を行うという意味では、やはり同類である。観察された諸統制変数を (説明変数の割当グループごとに) 縮約・バランスングする傾向スコア分析も、変数データの追加的採取を前提としたモデルである。

これに対して、分断デザインや自然実験では、説明変数への割当自体がランダムに生じ

ているケースを見出すという意味で、固定効果の追加的観察に依存することなく個体の異質性の影響を除去することを可能にしている。ただし、いずれの方法においても、自然実験状況が運良く生じている（あるいはそれをコストをかけて生じさせる）ことが必要になるため、適用できるケースは限られてくる。

2.3 混合効果モデルにおける異質性の扱い

上記いずれの手法においても、個体の異質性は除去すべきノイズであった⁶⁾。他方で、個体の異質性について、以上の各種手法と異なったアプローチをとっているのが、ランダム効果モデル (random-effects model) や混合効果モデル (mixed-effects model) である。両者の区別は必ずしも明確ではないが、前者が主にテキストブックにおいて、誤差調整という文脈で（実際には最尤推定が使われている場合でも）GLS 推定量と平行に位置づけられているのに対して、後者は単なる誤差の調整とは別の目的で用いられることが想定されている。

混合効果モデルは $y = X\beta + Zu + \varepsilon$ という基本形で表現できる。ここにおいて、 Z は説明変数であり、たいいてい X (切片を含む) の部分集合である。 u は個体 (クラスター) 間の変量効果ベクトルであり、個体 (クラスター) 数 $\times Z$ の行列である。ここで個体とは、データの最小単位が集合して作られる集まり (クラスター) を指す。生徒個人が最小単位であるとき、その集まりである学級などが想定される。

すでに述べたように、混合効果モデルは、変量効果最尤推定と基本的に同じものであるが、その主要な目的は異なっている。GLS 系の手法と互換的に用いられる変量効果最尤推定量においては、想定される機能は誤差の調整である。しばしばパネルデータを用いた分析において、固定効果系の推定量と変量効果推定量が有意に異ならないという条件が満たされる場合に、(たいいていの場合) 有効性の観点から変量効果推定量が選択される。OLS 推定量は、同一個体の反復観察に伴う誤差の拡大を調整する機能を持たないためである。

これに対して混合効果モデルにおいて目指されるのは、固定効果 (β) ではなく、変量効果 (u) の不偏推定である。変量効果推定量は、個体効果のひとつの推定方法である。すなわち、デモグラフィック変数として観察された分 (固定効果) を除いても残る個体の異質性そのものを推定するのである。固定効果の影響が十分に考慮 (除去) されている場合、 u は「最良線形不偏予測値 (Best Linear Unbiased Predictor: BLUP)」となることがわかっている。BLUP はしばしば経験ベイズ推定量、縮約推定量とも呼ばれ、単純な個体ダミーの影響の大きさは異なった値が算出される⁷⁾。

以上のように、混合効果モデルの発想は、個体の異質性の効果を除去した上で固定効果の推定値を得るという方針とは逆に、固定効果の影響を除去した上で個体特有の効果の大きさを追究する、という点にある。因果推論においては異質性がノイズであるのに対して、混合効果モデルにおいては固定効果こそがノイズなのである。

このような混合効果モデルの分析方針は、因果推論の枠組みとは大きく異なるものであるし、それどころか、もっと「緩い」因果志向の分析とも相容れないものである。さまざまな個体特性 (もっとも根本的なものは遺伝的特性であろう) の効果を除去した上で、たとえば自主学习といった個体の主体的選択あるいは外部からの介入 (例えば能力別クラス編成) の効果を知りたいというのがほとんどの社会科学における分析の目的であるからだ。学習や教授法といった「後天的」な選択・介入の効果を除去した上で、個体「本来」の能力を知りたいといった分析は、通常はなされない⁸⁾。

ところが、混合効果モデルは、マルチレベル分析というラベルとともに、社会学分野で広範に活用されるようになった。これは、上記のような混合効果モデルの本来の機能からすれば理解することが難しい現象であった（筒井 2017）。実際、因果推論志向が強い計量経済学の分野においては、混合効果モデルあるいはマルチレベル分析はほとんど無視されている。例外はすでに述べた変数効果最尤推定であるが、これは推定値が固定効果推定量と一致する場合にのみ採用される、いわば「控え」のモデルにすぎない。

2.4 因果推論におけるデモグラフィック変数

計量経済学においては「控え」の役割しかあたえられていないマルチレベル分析が社会学において重用された理由について適理的な説明を与えることは難しいが、消極的な理由として、デモグラフィックデータを重視する計量社会学の特性に起因する、という説明を考えることができる。回帰分析といった、因果推論志向が相対的に弱い手法においては、デモグラフィック変数は個体の異質性を統制するために投入されるとはいうものの、観察された変数自体の影響の大きさは推定される。しかしマッチングにしる傾向スコアにしる、あるいは固定効果推定にしる、因果推論志向が強い手法においては、個体の異質性は文字通り除去され、そもそもその大きさが推定されることはない。

固定効果推定においては、個体内平均化あるいは個体ダミー変数の投入により、観察期間中に変化しない個体効果によるバイアスは除去される⁹⁾。ルービン・モデルにおいては、個体の異質性は、観察されたデモグラフィック変数が集約された値（スコア）として表現されるため、個々の変数の相関の強さは推定されない。

以上、個体の異質性という概念を軸に、代表的な計量手法を整理した。固定効果のバイアスのない推定を目的とした手法の場合、古典的 최소二乗推定量においては、観察を通じた変数効果から固定効果への「移し替え」が主要戦略となる。固定効果推定量においては、個体間変動のキャンセルによる個体特有の効果の除去が戦略となる。操作変数推定量においては、説明変数の割当に影響しないと想定される要因を観察し、その要因と説明変数との共分散のみを用いて回帰を行うことが、異質性除去の戦略となる。

ルービン・モデルのひとつであるマッチング推定量においては、説明変数の割当グループごとの観察された変数の縮約・バランスが主要な戦略となる。これらいずれにおいても、理論的に関連すると想定されるデータの観察が重要となる。他方で自然実験や分断デザインにおいては、なかば偶発的に生じた割当無作為化の状況を積極的に拾い出すことが、異質な個体によるセレクションを無効にする戦略となる。

これらに対して、全く反対に固定効果を除去した上で、個体の異質性の推定を目指すのが混合効果モデルである。本来ならば、（生まれつきの）個体特有の影響ではなく、社会制度等、構築された環境によって変動する固定効果に関心があるはずの社会科学分野において、混合効果モデルが（マルチレベル分析として）普及した理由の一つは、異質性をノイズとして除去することを目指す因果推論系の手法と違い、混合効果モデルが個体（クラスター）内で変動しない個体の影響を推定結果から消去しないからだと推察できる。

3 介入と変数の概念連関

3.1 回帰分析を巡る二つの立場

前節では、混合効果モデルにおいては、観察された個体の異質性に観察されない個体

の異質性を、固定効果を持つものとして観察された変数の変量効果として付随させるため（上記式における Z が観察された個体特性、 u が観察されない異質性の効果である）、デモグラフィック要因をモデルから除去せず、その影響の大きさを推定することが、社会学における活用につながったということを論じた。

この分析の方向性は、計量社会学における古典的な回帰分析の使い方においても同様にみられる。シエの整理によれば、回帰分析については二つのアプローチがある。ひとつは「ガウスの」アプローチであり、もうひとつは「ゴルトンの」アプローチである。

ガウスのアプローチは「類型的思考 (typological thinking)」に属すると言われ、ある現象の変動を固定効果と変量効果に分割する際、変量効果を誤差（典型的には観察誤差）だとみなす立場である。他方、ゴルトンのアプローチは「集団的思考 (population thinking)」に属すると言われ、固定効果をシステムティックな多様性とみなし、変量効果をその他の個体内 (within) 変動としてとらえる立場である。別の言い方をすれば、変量効果を（いまだに）観察されていない固定効果の集積だと考えるのである。

これらのアプローチの違いは、実際上の分析においてはそれほど顕著な違いとしてあらわれるとは限らない。シエも、フリードマンの言葉を引用しつつ、この違いは「非常に微妙 (too subtle)」であり、「同じ推定手続きを用いる」ことに違いはないと述べている (Xie 2007: 145-46)。異なるのは解釈、背景となる仮定、そして研究目的である。

その意味では、強い意味での因果推論志向のモデルの台頭は、類型的アプローチと集団的アプローチとの違いを鮮明化したといえる。というのは、たとえばルービン・モデルにおいては、明示的に変量効果を縮約し、その影響を消去することが目指されるからである。回帰分析において二つの立場が曖昧に混在する段階は終わり、多かれ少なかれ、分析者が選び取るアプローチに沿った手続きを遂行することが期待される段階に入っているのである。

3.2 介入の意味

回帰分析における個体の異質性の位置づけについては、すでにいくつかの考察が存在することがわかった。他方で、因果推論における介入の位置づけについても、いくつかの重要な考察がなされてきた¹⁰⁾。

ただ、前節での個体の異質性の位置づけについての議論を踏まえると、介入概念について従来とは別様の議論を展開することもできる。シエのいうゴルトンのアプローチにおいては、攪乱項、あるいは変量効果は、観察されていない（すなわち固定効果としてモデルに組み込まれていない）異質性の「集積」として理解できる。因果推論において「ノイズ」となる個体の異質性とは、実際には集積された結果として観察される異質性ではなく、その中に含まれている特定の異質性であり、それこそが自己選択を誘引する要因だと考えられている。これに対してゴルトンのアプローチにおいては、異質性はノイズどころか、説明の源泉として捉えられるのであった。

計量社会学でしばしば行われているように、デモグラフィックな説明のために線形モデルを用いる場合、結果の記述は次のように行われる。まず、ログリニア・モデリングでしばしば行われるような交互作用をモデルに組み込む場合、それは調査対象集団の分割を通じた説明になる。たとえば「男性/女性」カテゴリーに「高学歴/低学歴」カテゴリーをかけあわせて、四つのグループごとのアウトカムを記述・解釈する、といったやり方である。交互作用分析が行われないとすれば、それは観察された統制変数を通じた、メインの説明変数の平均的な影響の大きさにしか関心がない場合か、あるいは観察された統制

変数の各値を通じてメインの説明変数の影響の大きさに変化がない場合か、のいずれかになる。複数のカテゴリ変数間の関連を記述するログリニア・モデリングが主に社会学分野で発達したことを考慮しても、計量社会学では平均的な影響の大きさよりもむしろ実際に交互作用が観察されるかどうかに関心が持たれる傾向が強い、と考えられる。

このような手続きにおけるモデル構築と推定においては、説明が集団の分割を通じて行われ、そしてその分割は（理想的には）交互作用が観察されなくなるまで行われる。このプロセスにおいて、特定の（もっとも関心が持たれる）変数の各値ごとに、観察されていない要因を含めてデモグラフィック特性が＜均質に＞分布しているのかどうかは問題にならない。他方で「ガウスの」アプローチ、あるいは因果推論においては、観察された異質性については傾向スコアやマッチングによって均質化を施し（いわゆる「バランシング」）、観察されない異質性については実験や自然実験を駆使して均質化を図ることになる。

このような分析の手続き上の相違は、当然、介入や統制の意味の違いにも現れる。ゴルトンの、すなわち計量社会学的なアプローチにおいては、異質性の理解と説明は、実際に社会に存在している（意味的に理解できる）概念連関に則ったかたちで行われる。少なくともその傾向が強い（筒井・前田 2017: 205）。そのため、統制とはあくまで、観察された要因で表現されるデモグラフィックな集団において、ある変数の影響力が変わらないことを示すためか、あるいは統制変数の集団間の異質性があつたとしてもそこには研究上の関心がないために、平均的な影響力の大きさのみが分かれば良い、と考えているか、どちらかであろう。

ここでは、「他の条件が同じならば」という条件もまた、概念連関にそったかたちで考えられている。したがって、ある要因の概念の意味が解釈できる範囲でしか、統制の作業は行われなことになる。

これに対して因果推論枠組においては、変数の概念理解を度外視した均質化が行われる。無作為割当やバランシングは、まさに社会の多様な集団を均質化する手続きであり、傾向スコアを特定の要因に関連する諸要因の縮約情報として積極的に用いるのでなければ、割当の傾向性に何らかの社会的理解が付与されることは想定されていない。

因果推論枠組みのこの方向性を追究する場合、特に「措置」として位置づけられる変数については、できるだけ他の要因との連関から自由であることが求められることになる。これは、数値の《分析上》そうであると同時に、《概念的》にも当てはまる。すなわち、「他の条件が同じならば」ということは、突き詰めていくと、措置変数が概念的にも実質的に、できるだけ「空洞化」することを求めることである。このことにより、一部の因果推論派の論者の立場からすれば、デモグラフィック変数の因果効果については、それをそもそも「論じる」ことが難しくなる¹¹⁾。

デモグラフィック変数が、少なくとも単純なかたちでは因果推論における「措置／介入」として考えにくくなるのは、さしあたりはそれが「動かない／動きにくい」からだと思えることができる。ただし、因果推論における介入の特徴は、それだけだと捉えられない、ということを経験しておきたい。というのは、しばしば注目される自己選択との違いが、個体内変動がないことだけでは説明できないからだ。より根本的な違いは、デモグラフィック変数が通常、介入変数と違って、《意味的な厚みのある概念を反映》した変数であるということにある。たとえば「学歴」について考えてみよう。学歴は、家庭的背景、能力、シグナル機能、アスピレーションなど、さまざまな概念との関連において有意味な要因として理解される。関連する要素（「他の条件」）をすべて剥ぎ取ってしまうと、もはや私たちは学歴という概念について理解することが難しくなる。というのは、パールの語

彙を使うと、ある選択を介入として捉えるということは、因果ダイアグラムにおける「切除」を意味しているからだ。

因果推論の枠内においても、いわゆるメカニズムの分析においては、措置の効果の有意な解釈が可能である。たとえば学歴を割当することは原理的には可能であるが、その結果の解釈においては、学歴と関連する概念（たとえば教育を通じて取得された能力やスキル）を考慮した上で、効果を及ぼすメカニズムの分析・解釈がなされることになる。他方で、ルービン・モデルを使用した場合、焦点を当てる要因が社会的なプロセスにおいてどのような重要性を持っているのかについては、全体的な推論や解釈を行うことはできなくなる。

4 観察データに基づいた計量社会学の可能性

社会学においては、自らの「位置づけ」についてひとときわ意識が強く、「社会学とはなにか」についての考察の蓄積も多い（近年の一例として盛山和夫（2011）、佐藤俊樹（2011）など）。他方で、「計量社会学とはなにか」ということについての考察は、それほど目立たない。例外は、シエや石田が記録している、ダンカンが「遺産」として残したとされる一連の知見（Xie 2007; Goodman 2007）や、最近ではゴールドソープによる精力的な発信であろう。ダンカンは異質性こそが社会学の研究対象であると主張し（石田 2012: 11）、異質性をノイズとして捉える因果推論とは距離を置いた。ゴールドソープは、（計量）社会学を *population science* として特徴づける。*population science* とは、統計データを通じて人口集団の規則性を確率論的に把握するものである¹²⁾。ゴールドソープによれば、統計学的に発見された集団の規則性（たとえば出生率の低下）をもたらした諸原因について考察する際に、ルービンの *potential outcomes* アプローチの活動度が下がる（Goldthorpe 2016: 107）。

筒井（2017）では、計量社会学とその他の近隣分野での計量研究との距離を、「比較」の水準の相違に求めた。ルービン・モデルにおいては均質な群の比較が目指されるが、異質性を説明に用いる傾向が強い社会学では、条件をそろえるという手続きが、（ログリニア・モデリングにおいて典型的に見られるように）全体の傾向性を簡潔に記述する、という目的で行われることが多い。

これらの3つの議論は、全体（*population*）が異質性から構成されること、その規則性を記述する際の比較水準が「強い」因果推論のモデルとは異なっていること、という意味でつながっている。

本論文ではこれらの議論に加えて、計量社会学が参照する対象世界の概念連関において因果推論図式における介入がどのような意味を持ちうるか、ということを論じた。ゴールドソープがいう集団の規則性の記述は、基本的に観察データを通じてしかなされ得ない。パールが言うように、介入は諸条件からの「切除」であり、介入の効果についての分析結果を結合するだけでは、社会の多様性や変化を記述し、解釈することは難しい。

観察データにおける変数は、それぞれ一定の「概念の厚み」を持っている。これと対局なのが実験における介入で、介入は文脈から切断された操作として、概念的には空洞であっても実施し、結果を観察することができる。もちろん、観察のレベルにおいても外生性のショックや制度的な断絶におけるように、文脈からの切断は生じる。むしろ、だからこそ因果推論ではこれらのショックや切断を自然実験として、介入の代わりに用いるのである。他方で、社会的なプロセスの記述や理解は、通常はこういった断絶を含まない、要

因間の関連を想定してなされる。その意味で、社会的プロセスの記述が社会学的な課題であるとすれば、それは観察データに基づいて行われるしかない。

ただし、要素間の連関と断絶とは、そのあいだにグラデーションを含んだものであり、ある要因の概念をどこまで分節化するのは、当該概念の一般的理解のみならず、研究課題に依存する。たとえば「学歴」と関連する要素（家庭的背景、アスピレーション、能力、シグナリング等々）を切り離していき、学歴の一般的な概念理解を掘り崩すに至ったとしても、介入プログラムとしてそれが有意味である場合には、問題がない。あるいは観察データの分析において、出身家庭要因を切り離した上で学歴による違いをみるといった場合、進学が自己選択であることを勘案しても、出身家庭要因とは独立した学歴の影響の大きさをある程度推察することができる。肝心なのは、因果の捉え方の水準にそった分析を行うことである。

【注】

- 1) ただし、第二版(Morgan & Winship 2015)では6個に増えており、グラフ理論への言及も多くなっている。
- 2) ルービン派の立場からの学説上の立場の整理については、Morgan & Winship (2015: 90-101) など参照。
- 3) たとえば Kennedy (2008: 382) では、「マッチング」ならびに「傾向スコア」を使った手法について、「応用計量経済学」の章で2～3段落ほどの記述を費やしているにすぎない。代表的なテキストである Wooldridge (2008) においては、ルービンの研究もパールの研究も、いずれも参照されていない。
- 4) もちろん、因果推論を<強く>志向したモデルとそうではないモデルとのあいだに明確な線引きをすることは難しい。伝統的な計量経済学のテキストでは、しばしば古典的の最小二乗回帰モデル (classical least squares regression) を出発点とし、それが満たすべき条件に対する違反に対応するための「メニュー」としてさまざまなモデルや推定方法が説明されてきた。ただ、ロバスト推定量や GLS、トービットなどの一部の制限従属変数モデルは、古典的モデルの因果推論上の問題に対する対応というわけではない。
- 5) 社会学において自然実験を活用した研究が存在しないわけではない。アメリカのジャーナルでは、再犯に対する居住地の効果を推定した Kirk (2009)、勤労所得控除制度を自然実験として用いて乳児の健康に対する親の貧困の影響を推定した Strully et al. (2010)、ヨーロッパでは能力別クラス編成の効果を推定した Bygren (2016) などがある。ただ、経済学分野における自然実験の台頭と比較すると、その目立たなさは歴然であろう。
- 6) 例外は、効果の異質性についての分析である。社会学分野での研究については、Xie et al. (2012) など参照。ただし、この場合の異質性は観察されたデータによる効果のばらつきのことを指している。
- 7) 個体ダミーの大きさを β_c 、全体平均を β_0 としたとき、信頼性係数を用いて β_c を β_0 に近づける（縮約させる）、という手続きがなされる。信頼性係数は、 u 、 ε の分散および個体（クラスター）数 (k) から計算される。信頼性係数を λ としたとき、 $\lambda = \sigma_u / (\sigma_u + \sigma_\varepsilon / k)$ である。そのうえで、変数効果 u は、 $u = \lambda \beta_c + (1 - \lambda) \beta_0$ として計算される (Hox 2002: 29)。したがって信頼性計数が大きいほど、変数効果は個体ダミーの係数に近づく。
- 8) BLUP 法が用いられるのは、遺伝的要因にのみ関心がある家畜育種学においてである。家畜育種の分野では、介入（交配）によって残される遺伝的要因のみが重要なのであり、生育環境などによって後天的に獲得された特性がノイズになるからだ。BLUP 法を体系化した C. R. ヘンダーソン自身が、まさに家畜育種の分野で活躍した研究者であった。他方で、社会科学において混合効果モデルを活用する方法も存在する。それは、複数想定しうる個体（クラスター）の水準のどこに、被説明変数の変動を説明する要因が隠れているのかを探る手続きである（筒井 2012）。ただ、その場合には変数効果の個数を制限する必要もある。切片を含む固定効果（の一部）に変数効果を付与したモデルにおいて、変数効果の個数を q としたとき、変数効果の共分散が最大 $(q + (q^2 - q) / 2)$ 個分設定可能である（非対角成分 q 個と対角成分の半分 $(q^2 - q) / 2$ 個）。つまり、変数効果の個数の設定と分散成分の制約条件の設定など、ある程度ケースバイケースの判断が必要となる。分散成分の制約パターンの選択については、モデル間での尤度比検定などで行うことが可能であるが、それでも探索目的でマルチレベル分析を用いる際に、分析方針の恣意性に対する留意は必要である。
- 9) 観察期間外に変動する要因の影響は固定できない。また、観察期間内に変動する要因については、観察できたものについてはマッチング等で均質化できるが、観察できなかったものについては異質性の影響が残る。

- 10) 因果ダイアグラムにおける「バックドア基準」ならびに「切除」の手続き (Pearl 2000=2009; 宮川 2004), 潜在アウトカム枠組みにおける SUTVA (Morgan & Winship 2007: 37-40)などを参照.
- 11) 因果推論派によるパス解析に対する批判的コメントを参照. たとえば, "I suggest that it is epistemological nonsense to talk about one trait of an individual causing or determining another trait of the individual." (Kempthorne 1978: 15) .
- 12) ただし, ゴールドソープの「集団的規則性 (population regularity)」の概念は, ケトレーの「平均人」のアイディアに近い. 現代の社会学は, ケトレーの時代に比べると, 計算機能力の向上やモデルの発達により, より洗練された (一般人では観察できない) 規則性の把握ができるようになった, という位置づけである (Goldthorpe 2016: 14-15). シエの整理によれば計量社会学の系統は, ケトレーの平均人概念から距離をとったゴルトンに由来する (Xie 2007: 143). この点を整理するためには, ゴールドソープの確率の捉え方を理解することが必要だが, 別稿の課題としたい.

【文献】

- Bygren, M., 2016, "Ability Grouping's Effects on Grades and the Attainment of Higher Education: A Natural Experiment," *Sociology of Education*, 89(2): 118-36.
- Goldthorpe, J. H., 2016, *Sociology as a Population Science*, Cambridge University Press.
- Goodman, L. A., 2007, "Otis Dudley Duncan, Quantitative Sociologist Par Excellence: Path Analysis, Loglinear Methods, and Rasch Models," *Research in Social Stratification and Mobility*, 25(2): 129-39.
- Heckman, J. J., 2001, "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture," *Journal of Political Economy*, 109(4): 673-748.
- Hox, J., 2002, *Multilevel Analysis: Techniques and Applications*, Lawrence Erlbaum Associates.
- 石田浩, 2012, 「社会科学における因果推論の可能性」『理論と方法』27(1): 1-18.
- Kempthorne, O., 1978, "Logical, Epistemological and Statistical Aspects of Nature-Nurture Data Interpretation," *Biometrics*, 34(1): 1-23.
- Kennedy, P., 2008, *A Guide to Econometrics [6th Edition]*, Blackwell Publishing.
- Kirk, D. S., 2009, "A Natural Experiment on Residential Change and Recidivism: Lessons from Hurricane Katrina," *American Sociological Review*, 74(3): 484-504.
- 宮川雅巳, 2004, 『統計的因果推論：回帰分析の新しい枠組み』朝倉書店.
- Morgan, S. L. & C. Winship, 2007, *Counterfactuals and Causal Inference: Methods and Principles for Social Research*, Cambridge University Press.
- Morgan, S. L. & C. Winship, 2015, *Counterfactuals and Causal Inference: Methods and Principles for Social Research [2nd edition]*, Cambridge University Press.
- Pearl, J., 2000, *Causality: Models, Reasoning, and Inference*, Cambridge University Press. (= 黒木学訳, 2009, 『統計的因果推論：モデル・推論・推測』共立出版.)
- Rubin, D. B., 1974, "Estimating Causal Effects of Treatment in Randomized and Nonrandomized Studies," *Journal of Educational Psychology*, 66(5): 688-701.
- Rubin, D. B., 2005, "Causal Inference Using Potential Outcomes: Design, Modelling, Decisions," *Journal of the American Statistical Association*, 100: 322-31.
- 佐藤俊樹, 2011, 『社会学の方法：その歴史と構造』ミネルヴァ書房.
- 盛山和夫, 2011, 『社会学とは何か』ミネルヴァ書房.
- Strully, K. W., D. H. Rehkopf, & Z. Xuan, 2010, "Effects of Prenatal Poverty on Infant Health: State Earned Income Tax Credits and Birth Weight," *American Sociological Review*, 75(4): 534-62.
- 筒井淳也, 2012, 「マルチレベル分析を有効活用するには」『社会と調査』9: 102-6.
- 筒井淳也, 2017, 「数字を使って何をするのか：計量社会学の行方」『現代思想』45(6): 162-77.
- 筒井淳也・前田泰樹, 2017, 『社会学入門：社会とのかかわり方』有斐閣.
- Wooldridge, J., 2008, *Introductory Econometrics: A Modern Approach (4th edition)*, South-Western.
- Xie, Y., 2007, "Otis Dudley Duncan's Legacy: The Demographic Approach to Quantitative Reasoning in Social Science," *Research in Social Stratification and Mobility*, 25(2): 141-56.
- Xie, Y., J. E. Brand, & B. Jann, 2012, "Estimating Heterogeneous Treatment Effects with Observational Data," *Sociological Methodology*, 42: 314-47.

Special Issue Article

Quantitative Sociology and Causal Inference:

Understanding Society using Observational Data

Junya Tsutsui (Ritsumeikan University)

Abstract

Although statistical causal inference has become one of the major methods in quantitative research, systematic discussions of the meaning and impacts of it on quantitative sociology has not yet been found. In this paper, models of causal inference and related estimation methods are organized via the concept of heterogeneity. Above this arrangement, it is discussed, by using the sociological application of the multilevel analysis as a demonstration, that quantitative sociology has a tendency to treat heterogeneity in a different way from other quantitative research fields. As the conclusion, it is argued that there is a substantive difference between the approach of causal inference, which uses intervention or discontinuity from social processes, and that of the quantitative sociology, which usually refers to ordinary conceptual association in explaining social processes.

Keywords

quantitative sociology, causal inference, observational data

筒井 淳也 (つつい じゅんや) . 立命館大学産業社会学部 教授. 〒603-8577 京都市北区等持院北町56-1. jtsutsui@gmail.com. 研究関心: 家族社会学, 計量社会学.