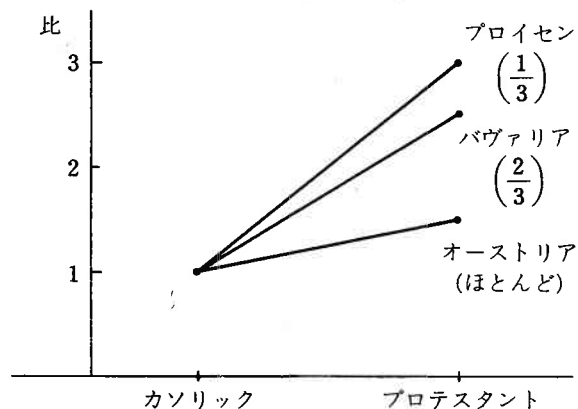


☆自殺率と宗教☆

デュルケムは『自殺論』の中で、個人の自殺傾向はその属する宗教によって異なり、一般にプロテスタントの方がカソリックよりも自殺傾向が高いことを示している。この現象に関する彼の中心的な論点は、「プロテスタントとカソリックのそれぞれにおける宗教的共同体の統合の度合いが自殺率に影響している」ということにある。これ自体が、「集団の統合度」という集合的変数の個人行動に対する影響を分析したものであるが、ただし、当然のことながら、自殺率のデータには「宗教的共同体の統合度」という変数が明示的に存在するわけではなく、この説明のデータによる検証はかなり間接的である。

一方、彼が示しているデータの中には別のタイプの集合的効果も見いだされる。ドイツ系の3つの地域、プロイセン、バヴァリア、およびオーストリアに関して、カソリックにおける自殺率を1としたときのプロテスタントの自殺率を図示してみると、図14-3のようになっている。



(注) 地域のあとの()は、その地域におけるカソリック人口のおおよその比率。

図14-3 ドイツ系の3つの地域におけるプロテスタントの自殺率 (カソリックの自殺率を1としたとき)

すなわち、カソリック住民の自殺率に対するプロテスタント住民の自殺率の相対的な大きさは、地域間で一定ではなく、地域におけるプロテスタント住民の占める比率に関連しており、後者が高いほど相対比率も高いという傾向が見られる。これは、自殺率に関して、プロテスタントかカソリックかという個人的説明変数(X)だけでなく、集団(地域)内におけるその分布(Xの比率)も効果を及ぼしているという例である。この現象は、「プロテスタントがその地域内で少数派であればあるほど宗教的な集団凝集性が高まって自殺傾向を小さくする」と解釈できる。デュルケムも、宗教的統合度の要因ほどには重視していないが、この「少数派要因」について気がついていて言及している。

社会調査法
放送大学教育振興会
1992

データから理論へ

データ分析とは、データをコンピュータに入力してアウトプットを取り出すことではない。それを通じて社会についての認識を深めることが目的であり、データを理論的に説明するという思考努力こそが、データ分析の技術を活かす道である。

1. 何のためのデータ分析か

データを分析することは、どんな実証科学にとっても、研究の一部を構成するにすぎない。

そもそも何のためにデータ分析を行うのか。それは、何か新しいことを知るためである。しかし、何でもが新しいことではありえない。当然ながら、すでに知っていることは新しいことではない。また、いままでまったく知らなかったことでも、たとえばランダムな数字の配列は知る意義のある新しいことではない。新しく知るといふことは、既存の知識と有意義な関連を維持しつつも、それとは異なるものを追加していくことであり、何らかの「発見」になるものである。

データ分析において、しばしば「あらかじめ仮説を考えておきなさい」と強調されるのはこのためである。「どんなことを知りたいか」という問題関心がまったく存在しない場合には、データ分析に着手することすら不可能である。データ分析の技術は、「データをどのように加工すれば、どのような側面が見えてくるか」を教えてくれるけれども、「どんな側面を見るべきか」は教えてくれない。したがって、データ分析の技術が深まれば深まるほど、問題関心をより鮮明で精密なものにしていくことが必要になる。さもないと、多くの手法を前にしてどれから取りかかるべきか途方に暮れるだけであ

る。

1章および9章で述べたように、新しく知ることは大きく分けて、「記述的に知ること=How」と「説明的に知ること=Why」の2つがある。このどちらのタイプにおいても「発見」はありうる。しかしどちらにおいてもそのためには努力が必要である。

まず、How について言えば、いかなるデータ分析の結果もそれ自体として何らかの「記述」を与えてくれる、ということに留意する必要がある。すなわち、何をやっても「記述」は得られる。重回帰分析をすれば、偏重回帰係数や決定係数が出てくる。クロス表を作成すれば、各カテゴリーのパーセントだけでなく χ^2 統計量などが出てくる。記述の材料に困ることはない。しかしまさにこの状況が「発見的な記述」を妨げる傾向がある。記述があまりにも容易であるために、その中から有意義なものを選び出したり探索するという知的な努力が怠られがちになるからである。すべてのデータ分析が自動的に発見をもたらすわけではない。「何を知りたいか」という問題意識を明確にし、「そのためには所与のデータについてどのような分析を行えばよいか」について既知の分析技術のノウハウの中から考えていかなければならない。

Why については、自動的にWhyに答えてくれる機械的なデータ分析法が存在しないために、容易なことではない。たとえば、「なぜ自民党支持率はかくも高いのか」という問いに対して、学歴や職業や所得などのクロス表をいくらつくっても簡単には答えは見つからない。このことは逆に言えば、もしもWhyの問いに対してある程度の答えの得られるデータ分析がなされたとしたら、それはそれ自体として十分に「発見」である、ということの意味している。すなわちWhyの場合には、Howとは逆に、その目的を達成することがより困難であるだけにそれだけ目的達成が「発見」を意味しているという傾向がある。

本章はデータ分析の技術をひととおり習得した人が、その技術を最大限に活かすためにはどのようなことをさ

らに注意しなければならないかについて、データと理論との関係を通じて説明するものである。

2. 理論とはどういうものか——『アメリカの兵士』の例

すでに1章において、理論とは「現象を説明する知識体系」であると述べた。ここには二つの特徴がある。一つは、理論は経験的に観測しうる現象に関わっているということである。逆に言えば、どんな理論も経験的世界との関係においてのみ意味を持つのであって、理論がそれ単独で意義を主張しうるものではない。第二には理論はそうした経験的現象を「説明」するものだというのである。すなわち、経験的現象と理論とは、説明-被説明の軸において「対象」と「メタ」との関係に立っている。このことは、理論そのものと経験的現象とは存在すべき地平を異にすることを意味している。

理論とデータとの関係を一つの有名な例で示してみよう。それはアメリカの社会学者たちが第二次世界大戦中に行った『アメリカの兵士』¹⁾の調査データと、それに関する「相対的価値剥奪の理論」および「準拠集団の理論」である。たとえば、データとして直接観測された内容は次のようなものである。

- ① 「あなたは本来召集を猶予されるべきだったと思いますか」という質問に対して、高校卒以上の高学歴者は11.6%がyesと答えたのに対して、それ未満の低学歴者は30.2%がyesだった。
- ② 同じ質問に対して、低学歴で20歳以上の既婚者は41%もがyesと答え、高学歴で20歳未満の独身者は10%だけだった。

これらがデータである。ここでまず注意すべきことは、このデータの中には一言も「相対的価値剥奪」とか「準拠集団」という概念は含まれていないということである。そればかりではなく「友人と比較して」というような言葉もまったくない。データは単にある質問に対する反応の仕方の比率が異なるカテゴリーの兵士たちの間で異なるということを示すのみである。

1) S. A. Stouffer, et al., 1949, *The American Soldier*, 2 Vols., Princeton University Press. また、ここで紹介した分析については、P. L. Kendall and P. F. Lazarsfeld, 1950. "Problems of Survey Analysis," in R. K. Merton and P. F. Lazarsfeld, eds. *Continuities in Social Research*, Free Press, を参照。

このようなデータを説明するものとして『アメリカの兵士』の著者たちや、R. K. マートン²⁾などによって提案されたのが「相対的価値剥奪の理論」であり、「準拠集団の理論」であった。これらは「人は自分の地位や状況を評価する際に、自分と似たような環境のもとにある人々の集団を想起し、その一般的傾向と比較して判断する」というものである。この一般的理論を上のようなデータに適用すると、「異なる集団の間で召集猶予に対する意見が異なるのは、それらの間で召集が猶予される人々の割合が異なり、それが高い集団からきている兵士ほど自分が猶予されなかったことに対する不満が高い」という説明が導かれる。(この説明を理解するためには、それら社会学者たちにとっては自明ではあったけれども、われわれにとっては必ずしもそうではない、当時の召集猶予に対するアメリカの政策を知っている必要がある。既婚者が未婚者よりも猶予される傾向があったことはわれわれにも理解できるとしても、高学歴者と低学歴者との関係は一見奇妙に見える。実態は、アメリカにおいては、低学歴の方が軍需産業における民間の労働者として必要とされていたため、高学歴者と比べると召集が猶予される傾向が高かったのである)。

この理論及びそれに基づく説明についても、まず注目すべきことは、これらはほとんどデータの中に直接観測されていない、ということである。「人々が自分と似たような環境のもとにある人々の集団を想起している」ことも「それと比較して判断している」こともまったく観測されていない。観測されているのは、あくまで上の①と②のようなデータだけである。

「実際に召集が猶予される傾向が異なる集団の間で異なる」という傾向は、当該のデータの中にはないけれども、当時の研究者たちにとっては常識的な知識の一部をなしていて、ある意味で観測された現象といえる。しかもこの点については、やや間接的ながら『アメリカの兵士』の中に次のようなデータがある。

③ 「あなたの友人や知人のうち、社会的に必要不可

2) R.K.Merton, 1957, *Social Theory and Social Structure*, Revised ed., Free Press. (森東吾他訳, 1961, 『社会理論と社会構造』みすず書房)

欠な民間の職業についているために召集を猶予された人がありますか」という質問に対して、高学歴者は19%がyesと答えたのに対して、低学歴者は79%もがyesと答えたのである。

したがって、学歴と召集猶予に対する意見との関係については、「友人・知人に召集猶予が多い学歴集団の方が、召集猶予されなかったことに対する不満が高い」という傾向は、データとして確認されている。しかしここでまだ注意して欲しいのは、いまデータで示されているのは、「低学歴の方が友人・知人に召集猶予が多い」ということと「低学歴の方が自分が召集猶予されなかったことへの不満が高い」ということだけであって、「前者が後者の原因である」とかましてや「そうした友人・知人を比較の基準にしている」というようなことは依然として観測されていないのである。どんな場合にもデータが自ら原因を語ることはありえない。われわれが手にしうるデータは常に変数の値の分布であって、変数の間に時間的な順序があるときでさえ、「何が何の原因か」までは指定してくれてはいないのである。

ただし、データ分析はこの点について、ある程度までさらに肉薄することができる。もしも準拠集団の理論が正しくて、「低学歴者の中で友人・知人に召集猶予が多い」ことが「低学歴者の中で自分が召集猶予されなかったことへの不満が高い」ことの原因であるとしたら、「単に友人・知人に召集猶予がいるかないかが不満に関係しているのであって、本来学歴とは無関係だ」と推察しうる。もしもこの推察がデータによって確認されるならば、それは「友人・知人に召集猶予が多いことが不満の原因だ」という説明、さらには一般的に相対的価値剥奪の理論あるいは準拠集団論を、立証はしないまでも支持し、その信頼を増すものであろう。

この推察に対応したデータ構造の挙示は、13章で述べたエラボレーションで行うことができる。すなわち、ケースをここにおける第3変数である「友人・知人に召集猶予者がいるかないか」で分割し、分割されたグルー

表15-1 友人・知人に猶予者あり-なしで分割されたクロス表

	友人・知人に猶予者あり			友人・知人に猶予者なし		
	召集に満足	召集に不満	N	召集に満足	召集に不満	N
高学歴	62.7%	37.3	335	94.4	5.6	1426
低学歴	63.3	36.7	1484	94.6	5.4	392

プごとに学歴と不満との関係を見るのである。この関係について、上記の推察は「学歴と不満とは独立に分布している」という仮説を主張していることになる。

分析結果は表15-1のとおりである。

この表は驚くほど仮説に合致している。友人・知人に猶予者がいるかないかでケースを分けると、召集されたことへの満足-不満に関して、学歴の違いはまったく消えてしまう。そして、学歴のいかにに関わりなく、友人・知人に猶予者がいる兵士たちの約37%が不満を表明し、そうした友人・知人を持たない兵士たちはわずか5%程度だけが不満を表明しているのである。したがって、推察は支持され、「単に友人・知人に召集猶予者がいるかないかが不満に関係しているのであって、本来学歴とは無関係だ」と判断することができる。

3. データ分析と理論との関係

探求のプロセス 理論とはデータを説明するものである。この基本的な関係は、上の『アメリカの兵士』の例でわかったことと思う。データと理論をつなぐ探求のプロセスを一般的に示すと以下のようになる。

- ① まず、説明されるべきデータが存在する。これは、現在分析しつつある調査データの一部であることもあるし、まったく無関係な別のデータであることもある。
- ② このデータを説明しようとする理論を考案する。これは一般には容易なことではない。いくら考えても、もっともらしい理論や説明的仮説が思い浮かばないことの方が多い。そのときは、それ以上の分析を断念す

るのみである。したがって、どんなに小さな思いつきでも、貴重である。

③ 何らかの理論が浮かんだら、次のような推論を行う。「もしも、この理論が正しいならば、現在手持ちの調査データにおいて、いかなるデータ上の構造が見いだされるはずであるか」。これが、理論をデータと関連づける接点をなす。上の例では、「もし、相対的価値剥奪説で説明できるなら、＜猶予の友人の有無＞でデータをコントロールすると、学歴と満足間との関連は縮小ないし消滅するはずだ」という予想がそれである。言うまでもなく、ここではデータ分析に関する知識がものを言う。どんな予想も、自分の知っている分析法の範囲内でしか思いつかれなければならないからである。

④ この予想に従ってデータ分析を行う。通常は、ここにおけるデータ分析が予想どおりに成功することはむしろ稀である。

⑤ 予想どおりではない結果が得られたとき考えられるのは、そもそも理論が間違っているか(②の誤り)、もしくは理論からデータ構造の予想を導く論理が間違っているか(③の誤り)のいずれかである。いずれであれ、予想どおりでない結果が得られた場合には、理論に立ちかえて考察しなければならない。

もう一度強調しておく、②の段階と③の段階とはまったく理論的な思考上の作業であって、データはその出発点であったり傍証であったりするけれども、その思考プロセスをつくり出すものではない。データが考えてくれるのではなく、研究者が考えなければならないのである。

理論=経験的 一般化の誤り

多くのテキストは、理論とデータ分析との関係についてたいへんな思い違いをしている。その典型的な例が、理論を経験的一般化と同一視する誤りである。

経験的一般化というのは、いくつかのデータを見ることによって、そこに共通に存在すると思われる一般的な傾向を抽出してくることである。たとえば、「40歳代の

☆「仮説」の3つの意味☆

調査法のテキストにはしばしば「仮説 Hypothesis」という用語が現れる。しかし、同一の用語でありながら、少なくとも3つの異なった意味で用いられている。

第一は、統計的検定における「帰無仮説」という用法である。これは、標本値から母集団の分布に関して考察を行う際に、母集団の分布を規定している母数の値について立てられる「仮説」であって、たとえば「全国の有権者の自民党支持率は実際に50%以上である」というような仮説である。

第二は、理論上の仮説である。これは、しばしば「公理」とか「前提」とも呼ばれる。ある理論を構成する、それ自体としては検証できない基本的な仮定であって、ニュートン力学の「三法則」などが典型的な例である。

第三は、データに関する仮説である。「アメリカの兵士」の例で、「徴兵を猶予された友人があるかないかでケースを分割すると、学歴と満足との間には関連はない」というのがこれに当たる。これは2章で述べた「作業仮説」に対応しており、データによって直接テストすることができる。

調査データの分析に最も関係するのはやはり第三のデータに関する仮説であるが、これが母集団に関する主張と見なされるときには、統計的な帰無仮説を用いた検定が必要となる。また、データに関する仮説も本来的には何らかの理論から導かれたものであることが望ましく、そのときには理論上の仮説とも関連することになる。

男性は美術鑑賞には行かない」、「おしぼりで顔をふく人は休日はごろ寝テレビでゴルフと野球をみる傾向がある」、「爪楊枝をくわえて食堂からでてくる人は、カラオケで昴とマイウエイを歌うのが好きだ」というようなデータがあったとする。これらから「動作のオジサン度」と「余暇活動のスタイル」とには関連があることが推察される。ここから「動作のオジサン度の高い人は、自己満足型余暇活動を行う傾向がある」というような一般命題を抽出するのが経験的一般化である。

① 多くの人々は、経験的一般化が理論の基本形態だと間違えて考えている。それは、「理論＝データの要約」という理解に基づいている。しかし、すでに何度も強調したように、「理論＝データを説明するもの」でなければならない。相対的価値剥奪説とアメリカの兵士のデータとの関係がこれを例証している。経験的一般化は、デー

タを要約的に記述することはできるけれども、データを説明していることにはならない。そのもととなった個々のデータがそれによって説明されるのではなく、データに含まれるいくつかの異なる概念の背後に、オジサン度のような「共通の要素」を想定し、その共通概念によってデータを記述しただけである。

経験的一般化は「概念の操作化」という考え方と密接に結びついている。そこでは、たとえば「オジサン度」という概念を「おしぼりで顔をふく」、「爪楊枝をくわえてでてくる」といった観察可能な諸指標から操作的に合成して定義することが試みられる。そうした操作的に定義された概念を用いて命題を立てることによって、命題のデータによる検証が可能になる。これが理論の検証だ、と考えるわけである。しかし、まず第一にこれはすでに述べたように何も説明しない理論でしかない。第二に、「オジサン度」というものの抽象的概念が操作的に正しく定義されているかどうかを判定する方法が、基本的には存在しない（「妥当性」の議論が多くのテキストには書かれているけれども、それらを読めばすぐにわかるとおり、「妥当性」の概念は定義できても、データによってそれを確認する手段は存在しない）。むしろ経験的一般化や操作的尺度の構成は、やっていけないことではない。それらは、有用な知見をもたらすこともある。しかし、それは説明されるべき素材としてであって、説明する理論としてではない。

4. 創造的データ分析

社会調査データの分析法をはじめて学びつつある人にとっては、データを説明する理論を新しく創出するということが、あまり身近な課題には感じられないかもしれない。実際、まずは分析法のさまざまな技術を習得することが、重要なことである。理論的に考える力そのものが、データ分析の方法に関する広範な知識なしには発展していかない。どのようなデータ分析を行えばどのようなデータの構造が抽出できるか、マスコミや官庁統計

で報告されたデータの背後にどのような社会的現実を読み取るか、あるいは、専門的な研究報告がどのようなデータ分析を用いていて、それから何をどのように解釈することができるか、そうした能力をまず身につけることが必要である。データを「読む」力がなければ、理論的に考えることもおぼつかない。

同時に、データ分析の技術の習得のときにも、それと並行して考えていかなければならないことがある。それは、データ分析の目的は何か、ということである。技術は目的があってはじめて意味を持つ。社会調査データに関しては、それは、分析を通じて社会についての理解を深める、ということであろう。それは単にデータをコンピュータに通してアウトプットをうるというだけではなく、データの背後にあって、データを生み出している社会的な現実について考えるということである。

研究とは、常に新しい、前人未踏の課題を解こうとするものであり、まだだれも説明することに成功していない現象を説明することは、まったく創造的な知的活動である。それはデータ分析の技術だけではなく、説明のための理論を考案し、それに従えばデータがどのようにうまく説明できるか、どのようなデータ構造が予測されるかを「考える」ことが必要不可欠である。データ分析の技術はそうした思考努力と結び合わされることによって、社会に関する理解の進展に寄与しうるものである。

(盛山和夫)

参考文献

1. 社会調査一般の参考書

Mann, P.H., *Methods of Sociological Enquiry*, 1968. (P. H. マン『社会調査を学ぶ人のために』中野正大訳, 世界思想社, 1982年) 社会学の学生向けに書かれた社会調査法の標準的な教科書。今日では珍しいパンチカードやソーターがでてくるが、1968年の刊行なのでしかたがない。

西平重喜『統計調査法(改訂版)』培風館, 1985年

世論調査の統計的基礎を解説した定評のある教科書。事例の紹介も行き届いている。

鈴木義一郎『例解標本調査論』実教出版, 1981年

標本抽出法の原理を、単純な母集団を前提に初学者にもわかるように説明している。

杉山明子『社会調査の基本』朝倉書店, 1984年

世論調査の実務的な知識を調査のプロセスに即して整理している。

安田三郎/原純輔『社会調査ハンドブック(第3版)』有斐閣双書, 1982年

データの収集から分析、さらには用語の解説まで、社会調査に必要な知識を詰め込んだ文字通りのハンドブック。

その他、以下のような参考書がある。

飽戸弘『社会調査ハンドブック』日本経済新聞社, 1987年

G. イーストホープ『社会調査方法史』川合隆男/霜野寿亮監訳, 慶応通信, 1982年

福武直『社会調査(補訂版)』岩波全書, 1984年

G. ギャラップ『ギャラップの世論調査入門』二木宏二訳, みき書房, 1976年

原純輔/海野道郎『社会調査演習』東京大学出版会, 1984年

林知己夫『調査の科学—社会調査の考え方と方法』講談社ブルーボックス, 1984年

宝月誠/中道實/田中滋/中野正大『社会調査』有斐閣Sシリーズ