

集計結果が分かれば良いのか

手集計を手掛かりに考える

日本社会学会で初めて発表したとき、司会者は西田春彦先生と安田三郎先生だった。その安田先生から誘われて大学院のゼミに参加し、その後、東京都目黒区と世田谷区の成人を母集団とした偏見についての標本調査を、先生の指導の下で行った。データ集計のことが話題になったのは、その準備中のことだったかと思う。1971年に刊行された『社会移動の研究』の内容を、先生は手集計カードとタイガー計算機を駆使して作成していた（手集計カードについては、福武直『社会調査』岩波書店、西平重喜『統計調査法』培風館などを参照）。

当時の一般的状況を私は知らないが、私自身は、社会工学専攻の院生時代に行った郵送調査以来、コンピュータを使って集計していた。そうは言っても、大学に一つしかない大型計算機センターが提供しているサブルーチンは科学技術計算用のものであり、世論調査タイプのデータには対応していない。そこで、機械工学科出身の助手・児玉文雄さん（現・東京大学名誉教授）の助言を受けて、FORTRAN言語の算術GOTO文を用いて単純集計プログラムを作るところから出発した。目の前にあるパソコンでSPSSなどの統計パッケージを使い、システムと対話しつつ数行のコマンドを作れば即座に集計結果が得られる、という近年の状況と比べれば、文字どおり隔世の感があるが、それでも、当時としては画期的なことだったと思う。1975年SSM調査で本格的に電算機利用が導入されるのは、その数年後のことである。

私がFORTRANを用いて手作りした集計や分析の方式は、今となっては古色蒼然たるものだが、電算機を用いて一挙に結果を出すやり方に対し、安田先生は必ずしも賛成でなかった。手集計も捨てがたい、というのである。単純集計にせよクロス集計にせよ、手集計カード上に記された回答カテゴリーを手掛かりにカードを配っていくとき、最終的に得られる分布に近い分布（カードの山）が初めから現れるとは限らない。それは途中でいろいろな姿を見せる。そのため、手集計でカードを一枚ずつ仕分ける作業をしながら、その時々分布が生み出される理由をおのずから考えることになる。そして、それは結果的に、調査対象が持ちうる潜在的可能性を考えることになる——それが安田先生の言い分であった。

全データを配り終わったときの結果は一つなのだからそれは無駄な努力だ、と評価することも可能だろう。しかし、基本に戻って考えてみると、標本調査によって私たちが得るデータは、仮に回収率100%だとしても、ありうるデータのつぎに過ぎない。ましてや、回収率7割というかつての回収率基準などほとんど望めない昨今である。私は、手集計を復活させよう、と主張しているのではない。私たちが実際に手にするデータには、調査票作成から実査を経てデータ入力に至る多くの過程でさまざまな歪みや誤差が混入してくる。そのことを自覚し、分析結果を柔軟に読み解く目を養うことが、分析手法がますますブラックボックス化する状況の下では、これまで以上に重要になると思うのだ。

海野道郎

東北大学 名誉教授