

「科学的」世論調査の価値<sup>1</sup>

—歴史と理論と実践の三位一体—

吉野 諒 三\*

“Scientific” public opinion poll as an achievement of Japanese statisticians

Ryozo Yoshino\*

本論文では、戦後日本の統計学の展開を、特に統計的標本抽出理論に基づく科学的世論調査研究の実践的方法論確立の観点から振り返る。統計学会75年の歴史のなかで、実のある統計学の発展は「歴史」と「理論」と「実践」の相互関係の中ではじめて必然となることを、戦後民主主義発展の基盤となった統計科学的世論調査研究が例証していることを再確認する。また一方で、近年、個人情報保護等の影響で調査環境が悪化し、調査遂行上の著しい困難がもたらされ、世論調査の統計的基盤が揺らいでいる。本論文では、この問題解決への試案を示唆するが、これが一つの契機となり、多くの統計学者の尽力により、近い将来、再び統計科学的世論調査が揺るぎのないものに再構築されることを期待する。

The main objective of this paper is to review the history of “scientific” methodology of public opinion poll developed by the Japanese statisticians from the post-WWII period downward. Here the term “scientific” poll means that it is based on the statistical random sampling theory which yields the estimation of sampling errors. This review exemplifies the importance of interrelationships between historical background, theory and practice. This may be closely linked to the later development of statistical philosophy, such as “tohkei-suuri (statistical mathematics)”, “suryoka (quantification theory)”, “multidimensional data analysis”, and “Science of Data.” Nowadays the methodology of public opinion poll suffers from low valid questionnaire returns because of so many rejections and frequent absences of potential respondents. I suggest some possible ideas to approach towards the problem. And I hope that the Japanese statisticians will develop a new practical theory to recover the credit of scientific public opinion poll in the near future.

*Key Words and Phrases:* CULMAN, postwar democracy, Hayashi’s quantification method, public opinion poll, sampling theory, Science of Data, Statistical Mathematics.

## 1. 戦後の「歴史」

すべての学問は、「歴史」と「理論」と「実践」の相互関係を離れて成功することはない。中国では古来学問は政治に直結しており、欧米では近代以降の学問は産業革命後の産業発展のためになされてきた。日本では、政治から分離し、特定の価値観から独立したものとしての学問が強調される時期もあったが、やはり現実には社会の課題解決へ結びつかぬ机上の空論が広く社会に影響を与えることはなかった。

本論文は通常の論文の形式とは離れ、多少、饒舌になろうが、戦後日本の統計学者が創り上

<sup>1</sup> 本論文は、輿論科学協会の発刊している「市場調査」、259, 4-13. (2004) に発表した「科学的世論調査法の価値」を、同協会の許可を得て、その後の社会調査環境の状況と本特集号の趣旨を勘案して加筆修正したものである。

\* 統計数理研究所：〒106-8569 東京都港区南麻布 4-6-7

げた世論調査の方法論の統計学的基盤の意義を再考する。日本統計学会が75周年を迎えた今日、これが近い将来の統計学のさらなる発展の一助になれば幸いである。

近年、「データの科学」という言葉が各所で使われ始めているが、筆者は「データの科学」という言葉を2001年頃より、故・林知己夫 第17代会長(1985-86年)が中心となり唱えてきた意味で用いている(朝倉書店・データの科学シリーズ, 林, 2001; 吉野, 2001)。すなわち、現実社会の課題解決のための調査や観測の企画から始まり、データ収集からデータ解析までの全過程を通して現象を理解するという哲学である。戦後の数十年にわたり、数理統計の机上の論理に反発して登場した「統計数理」(水野, 2003)(これは研究所の名前ではない)、「数量化」(林・水野・青山, 1952)(これは特定の多変量解析の技法のことではない)、「行動計量学」(林, 1993b)、「多次元データ解析」(多変量解析と数学的には強く関連するが、哲学的には区別されるべきもの)(林, 1993b; 林, 2001, pp. 107-108; 林・鮑戸, 1976 参照), そして「調査の科学」(林, 1984)の延長上に昇華された、人文社会科学分野の複雑現象を解析するための新たな統計哲学、それが「データの科学」なのである(この経緯については、林, 2001; 高倉, 2003; 高橋, 2006; 森本, 2005 も参照)。

この一連の統計哲学の流れは、特に「戦後民主主義発展の必要性」と統計学とが邂逅した統計的世論調査の歴史と重なる。

日本において、統計的無作為標本抽出理論に基づく本格的な実践的調査は、昭和23年の「日本人の読み書き能力」調査に始まる(読み書き能力調査委員会, 1951)。日本語の漢字使用が初等・中級教育に困難を生じ民主主義の発展を阻害しているのではないか、日本語をローマ字化すべきではないかという議論があり、米国から教育視察団が訪れている。しかし、当時の占領軍は彼らの考えを直ちに押し付けるのではなく、現地の実態と住民の意向を尊重して事を進めるという方針をとった。それが民主主義の大義であった。文部省の統括の下に、日本全国から各分野の関連研究者が集合し、戦後初めての学際的研究調査プロジェクトが、現実の政策立案のための基礎情報収集として開始された。当時この仕事に携わった人々はCIE(民間情報教育局)から入手した書籍を勉強しながら、日本の事情に則した標本抽出計画を練った。

実は、標本抽出理論自体は戦前、既に1928年の亀田豊治郎の先駆的研究により、単純ランダムサンプリングの精度計算がなされ、日本における第一回のISI(1930年)で報告されていた。(翌1931年、日本統計学会設立。)亀田は1920年の国勢調査に参加し、また1923年の関東大震災の時、抽出標本から推計値を求め、速報していた(高橋, 2004, p. 22)。しかし、これは当時の社会情勢のためか、世論調査に関しては本格的な実践には結びつかず、戦後に米国から輸入されるという形になってしまった。どんなにすばらしい理論でも、広く世の中の必要性、実践に結びつかねば発展しないという例の一つである。

戦後の読み書き能力調査においては、実践的検討のために、小田原市で全数調査を遂行し、その有効回収票の一部を無作為抽出し、推定値と実際の値とを比較し、標本抽出理論(推定誤差の計算)の有効性を確認したという。全国調査結果の評価については、はじめは学者の中で意見が割れたが、統計学者の解析により日本人の読み書き能力は十分に高いと確認され、結局、「日本語」は救われたのであった。

この十数年、グローバリゼーションと称した浅薄なアメリカナイゼーションが、経済システムのみならず教育システムにまで影響を与えてきたが、世界各所での悪弊が報告され、米国自体も方針を変更してきた。われわれは過去の貴重な経験に鑑み、国家の政策を統計的基盤に立つ実証的証拠に基づいて慎重に検討しなければならない。

読み書き能力調査の成果を受けて、戦後民主主義発展の基盤となる科学的(統計的)世論調査が官民学で推進されていくことになった。「老若男女、皆、一人一票の民主主義」の統計的

基盤が確立されていったのである。ここで開発された実践的標本抽出方法を活用し、一方では政治、経済、産業の広範な分野で各種の統計的調査の実践的方法が開発され、また他方で昭和28年には統計数理研究所の「日本人の国民性」調査が開始され、やがて昭和46年以降の「意識の国際比較調査」(林他, 1998; 林, 2000; Yoshino, 2002; 吉野, 1994, 2005a; 吉野編, 2007)へと発展することになる。この研究は、米国の General Social Survey (GSS), Michigan 大学の Inglehart らの World Value Survey (WVS), EU 委員会による EuroBarometer など、諸外国での同様の時系的調査が開始される契機となった。昭和20年代の日本は各種の調査研究が花開いた時代であったと言われる。

これらは表の話である。現在の視点で当時の世界の流れを眺めると、戦前、戦中、戦後にわたり、米国社会学会では戦争相手国の「国民性」や「読み書き能力」が戦略や占領政策立案のための大きなテーマとなり、各国、各地で関連する調査研究が遂行されていたことが分かる (Inkeles, 1996, Ch. 1)。例えば、戦中の太平洋諸島では、日系人による日本人捕虜の尋問調査により戦略が大幅に効率化され、戦後の占領政策作成のための重要な情報を収集することに成功している。山本 (2001) の著書「日本人捕虜は何をしやべったか」の表 (p. 181) には天皇、軍部、マスコミ、政府に対する信頼感のデータが示されている。天皇への不信は0%, 軍部への不信の高さ、マスコミへの信と不信の拮抗など、わずかこれだけの限られたデータだが、戦後の天皇の位置、軍部の処置、政府やマスコミの改革など、占領政策立案に的確に反映されたようである。

戦後には各地や各国で、民主主義発展の必要条件としての読み書き能力 (一部のリーダーに盲目的服従をしない条件としての学力) が調査されたが、日本調査の背景には、1932年のトルコの国字改革に従事した経験を念頭に置き、検閲の簡便化をも考慮していた Hall が関与していたことがある (アンガー, 2001)。しかし、当時の彼らが卓見であったのは、調査遂行を信頼できる現地の研究者 (統計学者、文化人類学者) に任せたことである。調査主体 (占領軍側) が本当の目的を日本側には知らせなかったために、日本人の現地調査監督者や調査員にバイアスのかからない調査 (一種の DOUBLE BLIND 法) ができ、結果として、現地の実態や住民の意向を的確に調査できたといわれている。それらの調査に関与した人も、当時すぐには自身が関与した仕事の本当の意味 (米側の目的) は知らず、後になって気がついたらしい。現実には、米国内部でも立場により日本語のローマ字化には様々な議論や思惑があったようだが、形としては先の「読み書き能力調査」の結果を受けて、日本人の十分な能力を確認し、日本語は生き残ることになったのである。

他方で、戦後、米軍のアジア戦略上の要地であった奄美大島、琉球などでも住民調査が遂行され、Hall から米国オハイオ大学の日系人 Ishino 氏に極秘の奄美調査報告書が送られている。その報告書や手紙は POSR (Division of Public Opinion Survey Research) 文書として数十年ほど前に機密解除後、日本に返還され、現在は (財)日本世論調査協会が保管している。この Hall の手紙の中には、「現地の優秀な若手の統計学者 (数年前に亡くなられた) 水野坦氏と2人の文化人類学者の助力を得て調査を遂行した」という主旨の記述がある。現実には、住民運動や政府間の裏交渉など複雑な政治上の駆け引きがあったのだが (エルドリッジ, 2003), いずれにせよ、形としては住民調査の結果を受けて奄美は戦後比較的早い時期に日本へ返還されたが、沖縄の返還は昭和47年まで延ばされるという運命をたどることになる。

今回のイラク戦争開始前に、日本の識者が、米軍による「日本の戦後占領モデル」は適用可能かと問われ、多くは「明治以降、民主主義の歴史のあった日本と中東諸国とは異なる」と回答したそうである。しかし、戦前、戦中の慎重な調査研究、戦後には、信頼のできる現地人によるバイアスのない調査を遂行し、それを施策の基盤とした「米国モデル」の重要性を米国自

身が十分に認識していれば、今回のイラクでもはるかに適切な対応ができたであろう。今回は、失敗を続けた後、戦後の「米国モデル」に近い方針へ進み始めたようだが、初期の方針の誤謬による混迷はいまだに深い。

## 2. 「科学的」な世論調査

放送法には、(NHKの放送局は)「放送及びその受信の進歩発達に必要な調査研究を行なうこと」(9条1項2号)とともに、「公衆の要望を知るため、定期的に科学的な世論調査をおこない、かつ、その結果を公表しなければならない」(44条2項)と定められている。ここでのキー・ワードは「科学的な世論調査」である。林(1977)によれば、「科学的」とは「精度が計算可能という意味」である。そして、「科学的調査」は、当然、科学技術の進歩とともに変わるであろうとされた。

確かに科学技術は進歩し、電話やコンピュータネットの時代になった。しかし、現在の日本の調査は、むしろこれらによって非科学的になっている。一人一票の民主主義を尊重し、日本人の成人(有権者)全体を母集団として、統計的無作為標本抽出により母集団の真の値を推定誤差とともに計算する方法から、コストを理由に、安易な方向へ流れてきてしまった。

戦後、世論調査方法が確立し始めてから、実は日本の調査関係者は米国の多くの調査はクォータ法(性別・年齢別などあらかじめ指定した属性のみについて勘案し、人口分布比例で回答者を集める)であり、完全な統計的ランダムサンプリング(理論的にはあらゆる属性について恣意的なバイアスのない回答者の標本収集)ではないことに気づいた。過去の米国の大統領選挙等の幾度かの予測失敗は、クォータ法の欠陥が一因であったといわれる。つまり、世論調査の方法論に関しては日本の方が余程、科学的で民主主義的であるといえる。しかし、その誇りを科学技術の浅薄な応用が崩す危惧がある。

例えば、昨今、RDD方式の電話調査が盛んに用いられるようになってしまったが、RDDでアクセス可能な母集団は明確ではなく(明らかに有権者全体ではない)、「回収率」が計算できず「回収数」表記で済ますというのは、既に「科学的世論調査」を放棄し、戦後育成した民主主義の基盤をマスコミ自身が蔑ろにしてしまったことになる。某新聞社では、90年代初めの電話調査の試行の時には、面接法と電話法における調査モードの差(方法による回答差)及び標本抽出方法の差に関して慎重に配慮されていたが、ある時点から双方の差違を無視したグラフが表われ、やがてすべて電話法(住民票からのサンプリング調査もRDDも峻別せず)に切り替えるという流れを作ってしまった。(注. 80年代末に日本全世帯の電話保有率と電話番号の電話帳登録率ともに十分に高くなったことがあったが、そのすぐ後には携帯電話の急速な普及に伴う世帯固定電話の保有率低下および電話帳登録率の著しい低下があった。)

インターネットの使用者から多数の回答を求めるという「インターネットアンケート」をも「インターネット調査」あるいは「ウェブ調査」と称するようになってきている。我々、調査研究者は、長年、統計的無作為標本抽出理論に基づくものを「調査」または「標本調査」とし、そうではないものを「アンケート」として峻別してきた。この区別は、自分たちが統計学を理解している調査研究者であることの誇りであった。しかし、従来の調査の発展としての「インターネット調査」ではなく、統計学には無知で、民主主義の理念にも鈍感なエンジニアが多様なITビジネスのひとつのアイデアとして、それを唄うようになってしまっているのである(本田・本川, 2005)。(注. 例外的に厳密な統計的標本抽出法による「インターネット調査」を施行している機関があるが、抽出標本へのPCやインターネット接続環境の確保等、費用は通常の面接法よりもかなり高くなる。)

勿論、「それらの方法では情報が得られないから、調査をするな」というのではない。即時

性のある情報を得るには重要な手段の一つであろうから、慎重な研究とともにその特性を生かして、「世論の動向」を探るにせよ、「市場調査」にせよ、有効に活用すべきだが、それを安易に「世論調査」とは呼ぶべきではない。「世論」をただ集めるのが「世論調査」なのではなく、理論と歴史的背景に留意すべきである。

1980年代より、マスコミでも電話調査の可能性が慎重に研究され始めていたが、その研究が十分に尽くされていない段階で、なし崩しに「世論調査」として電話調査が広く使われるようになってしまった。この理由の主要な一つとして、面接調査における拒否率・不在率の増加、したがって低回収率が問題とされた。しかし、これは「精度」に関わる問題であって、「科学性」が損なわれる問題ではない。回収率が分かれば、非回収層の分布（未知）をも考慮に入れた計算は可能である（吉野，2006）。精度が悪い（誤差巾が大きい）だけのことである。しかし、現今の電話調査の現状を見ると、結局は回収率向上にも繋がらず、特にRDDでは必ずしも回収率計算すらできないことが多い（注。一般世帯の登録番号集団の分布を予備調査である程度捉える努力がなされている機関もあるが、その場合も推定回収率が著しく改善されているわけではない）。回収率の代わりに「回収数」を明記するというのでは、誤解を助長する。現在の電話による調査では、多くの場合、先人たちが定義した科学的「世論調査」はなされてはいない。意見が回収されようがされまいが、そもそも原理的に母集団全体にアクセスすることのない方法では、「手続きとしての民主主義」、「一人一票の民主主義」を守る世論調査ではない。電話やコンピュータ・ネットという科学技術の発展が、科学的世論調査を消滅させようとしているかのようである。

否、これはそれらを扱う人間の問題である。繰り返しになるが、一般的な意味での世論の動向を知るために電話やインターネットを用いるなどというのではない。市場調査や一般社会調査の道具としては価値ある発明であろう。しかし、今のほとんどの電話調査等は戦後の世論調査の方法論の確立期に調査関係者によって合意され、定義された「世論調査」ではないといっているのである。「世論」は日常用語でも使われるが、特に日本の場合は、それは統計学的標本抽出理論に沿って厳密に実践的方法論が発展し、その操作的定義に則ることが合意された歴史の上にあることを知らねばならない。我々の歴史の中に生まれた「科学的世論調査」の意義を慎重に再考する必要がある。

世論調査は英語では public opinion poll であるが、poll は「投票」の意味でもある。投票率が低かったからといって選挙が無効になることは、普通はない。「あなたのご意見を政策立案に反映するための調査である」と明言しながらも、なお「回答拒否」であるならばそれも回答の一種、世論の一側面と考えられないわけではない。したがって、奇妙な方法に流れるよりは、結果として低回収率であったとしても、全うな科学的世論調査の方法に則った結果のほうが、たとえバイアスがあろうと確実な情報が引き出せるという考え方もありえる。バイアスがあるデータが悪いのではなく、バイアスがどちらにどれだけあるかも想像できないようなデータのほうがよほど厄介であろう。

一方、ESOMAR (European Society for Opinion & Marketing Research) の倫理規定には、選挙予測調査結果の公表における要件として、予測に耐えうる「十分なサンプル数」がある（小林，2003）。しかし、現実に日本で行なわれている各選挙区（統計的には無限母集団）で500程度のサンプルサイズでは、仮に理想的な単純無作為抽出による回収率100%の有効回収サンプルであったとしても、標本抽出誤差は小さくはない。さらに非標本誤差を加味すれば、大きな誤差となり得る。そういう誤差を超えた大差があれば予測も当たろうが、本当の問題は、僅差をいかに上手に判別し、予測するかであろう。現実には、選挙予測のための世論調査に限定すれば、「実際に投票する人の母集団」は世論調査の母集団とは異なるので、世論調査結果を

各属性集団の重み付けなどを考慮した予測モデルに組み入れて解析することが多く、その当否は予測のあたりはずれという明確で厳しい基準で判断される。通常の世論調査は、そのようには直接、検証することができないので、操作的方法の厳密性を強調するのである。

また、厄介なのが、拒否率の増加、在宅率の減少などによる低回収率の調査データに対して、回収標本と母集団との属性分布の差を比較して、「重み付け補正」しようという、現場を知らぬ机上の論理を振り回す研究者の誤謬である。最近、アメリカの影響を受けて、傾向スコアを利用しようとするものもみられるが、狭義の「世論調査」に限定して言えば、それも同じ誤謬を背負っている。アメリカのようにもともと標本抽出法の観点からは質の悪い調査データには、実験心理学の極端に限定された小標本を扱うときのように数字を加工する研究が多くなりがちである。その是非は、この数年のアメリカ大統領選や中間選挙に関わる世論調査に基づく選挙結果予測の失敗を眺めてみれば明瞭であるが、米国なりの歴史的、政治的背景を考慮すると、それもやむをえないのかもしれない。(選挙予測ではなく)世論調査における「重み付け補正」の誤謬については、林(2002, p. 66)、林・田中(1996)や吉野(2002b)に詳述してある。(端的に言えば、回収層のデータに対して各種の統計的「補正」をしてもわずか数%の差がでるだけであるが、回収層と直接観測不能な未回収層の意見分布の差の効果の影響の方がはるかに大きいことが知られている。したがって、直接検証不能な統計的仮定の下での「補正」より、まず、調査現場で質の高いデータ収集の努力すべきという常識が尊重されている。)

戦後の面接調査研究では、標本抽出誤差の理論のみならず、調査員の不正、データの偽造、回答者のウソの率、データ入力 of 誤謬の率など詳細に「非」標本抽出誤差についてもさまざまな形の実験調査研究が年月をかけてなされ、調査方法が確立していったことを再認識しよう(林, 1984)。

### 3. 「実践」のために

#### 3.1 科学的な「測定」の理論

これまで記述したことは、「道理」に立った言葉である。道理を前面に出されると人は黙するしかない。しかし、現実(本音)を無視した道理に人々がついて来ることはない。この2、3年ほどの国勢調査の協力拒否の増加、日銀や内閣府の調査における民間調査会社の不正は、この戦後の数十年にわたる調査回収率の低下、さらに個人情報保護法成立後の自治体の住民基本台帳や選挙人名簿の閲覧拒否の行き過ぎ、調査を語った凶悪犯罪の増加のための住民の調査協力拒否の極端な増加など、問題の深さを再認識させた。

我々は、現実の困難な調査条件の下で、どこまでその「道理」に応じられるのかを柔軟に考える必要がある。

筆者はこの数年、前節で述べたような批判を講演や論文等で繰り返してきた。現在では、マスコミの世論調査部の報道記事では、従来の住民基本台帳等を用いた統計的無作為標本抽出による面接調査は「世論調査」、RDDなど電話による調査を「電話世論調査」と名称を区別し報告されている。またRDD調査でも、電話会社が公表する情報が限定されている下で、日本人の成人全体(母集団)と固定電話でアクセス可能な集団(世帯、または個人)との対応関係を研究する努力も真摯に続けられている。しかし、いずれにせよ、著しい低回収率のもとでは、回収層と未回収層との分布の差異の情報を乗り越える統計理論はない(統計学は隠れた情報を浮かび上がらせるのに活用できるが、そこに本当に存在しない情報は発見できない)(林・田中, 1996)。

筆者は近年、CULMANと称する国際比較研究のパラダイムを展開している(吉野, 2005a)。これは一般社会意識調査のためであり、「世論調査」ではないが、その基盤となる標本抽出方

法の検討という意味もある。日本以外では、標本抽出における推定誤差理論の教科書が想定するランダムサンプルをとることは著しく難しい。欧米では、目標標本数を回収するまでランダムウォークを続ける方法や、性別・年齢層や人種など、あらかじめ指定した属性のみに着目して目標標本数を割り当て、該当する回答者を求めるクォータ法などが用いられることも多い。また、前節で日本の電話調査法を批判したが、例えば米国 GALLUP 社の標本抽出担当研究者（吉野，2007年3月インタビュー）によると米国では携帯電話と固定電話の番号の混用や「電話による市場調査」禁止法の影響等の下で、回収データのバイアスを補正するためのウェイト計算の非一貫性（性別・年齢層・人種比率をセンサスデータにあわせてウェイトをかけるが、あまりに大きなウェイト比のときは想定外のバイアスの助長を恐れ、トリミングする等々）も見られ、実態は混沌とした数字いじりに苦心している（吉野，2007）。

このような状況では標本抽出による推定誤差は計算できず、統計学的には望ましくないとされる。だが、多民族社会、階級社会あるいはその名残りのある社会では、何よりもまず、民主主義の公平として、それらの属性についての公平性が求められる。統計的科学性は、その後の問題となる。我々の国際比較調査でも、例えば近年の中国調査では、住民基本台帳等の名簿が活用できない場合のエリアサンプリング（地点は人口比例抽出、各地点では住宅地図を作成し世帯抽出、さらに Kish 法で個人抽出）を試行しているが、著しく低い回収率の問題は解決できていない（吉野編，2006，2007）。日本で同様のエリアサンプリングの試行した例も報告されているが、回収率 10 数%となり結果の公表に至っていない（鈴木，2006）。

我々は、現実の困難な調査条件の下で、むしろ、今一度、測定論の基礎に戻ってこの問題を考察する必要がある。以下は、試行錯誤の羅列に過ぎないが、真摯な統計学者が厳密な理論を構築する契機の一つになればと思い、その概略を記すものである。

公理的測定論（Narens，1985；吉野，1989，2007）において、「有意味な」（meaningful という専門用語）統計量とは「許容される尺度変換の下での不変量」と定義される。例えば、ある人の身長が 170 cm という場合の「170」は cm 単位のものさしで測った結果であり、同じ身長を inch 単位で測れば別の数値を与えるので、長さの尺度を cm でも inch でも測ることが許容される条件の下では「有意味」ではない。しかし、複数の人々の身長を cm で測って大小比較した「順位」と inch で測った「順位」とは一致するので、この順位は「有意味」な量である。つまり、許容される尺度変換（cm でも inch でも）下で、不変な量のみが科学的な測定量としての必要条件を満たす。

例えば、マスコミ各社の調査での内閣支持率は数値としては完全に一致するとは限らない。各社の調査方法が固定されている限りは、各社単独では有意味ではあるが、異なる時期の異なる社の結果をそのまま比較するのは有意味ではない。しかし、各社毎の経時的データの上昇下降トレンドは経験的にはほぼ一貫しているという意味で、そのトレンドは、各社共通に有意味ということになる。この有意味性を、世論調査の解釈や定義に持ち込むことを考慮しよう。これにより、狭義の科学的世論調査、すなわち「母集団からのバイアスのない標本抽出の厳密さ」にのみ依拠する「科学性」を乗り越える手がかりを模索してみよう。

上記の視点からは、操作的定義が明確でない調査の結果は、その不変性が議論できない。RDD の多くは、電話番号集団と国民（有権者全体）の集合との対応関係が明確ではない限り、そして統計的無作為抽出を望む限りにおいては、その結果には疑義がある。しかし、その対応が不明でもその結果が安定していると考えられる限りは、「測定量」としては有意味となる。これは短期的には成立するだろうが、（固定および携帯）電話所有者集団の変動が大きいときは成り立たない。また、面接調査による内閣支持率の上昇下降トレンドよりも電話調査によるトレンドのほうが急峻となる傾向が報告されているが、「トレンドの方向」は面接と電話のモー

ドを超えて有意味となる場合が多い。あるいは、時系列的研究で、電話の結果が面接の結果をあたかも先行するようなことも判明するかもしれないが、その場合は時間の幅（時間遅れ）を考慮に入れた一致性（有意味性）が検討されよう。

このように、許容される調査方法の範囲（集合）の中での有意味性を考えると、その「範囲」と「有意味な量の集合」とには相補性があり、一般に許容範囲が広いと有意味な量の種類は少なくなる。しかし、その制限の中で、狭義の理想的標本調査と同じ土俵で RDD 調査やインターネット・アンケートをも論じることは可能であろう。（ただし、近年の内閣府の比較実験調査では、面接や電話調査の母集団と、インターネット・アンケートでは著しく母集団が異なり、後者では家庭生活や社会への不満が断然と高く出ている。）

次節では、「(科学的) 世論調査」と「世論動向を知るための基礎データ収集」との区別を念頭において、さらに具体的に述べよう。

### 3.2 パターン構造を読む

1970～80年代、ISI 会長 Tukey (1977) は探索的データ解析を唱えた。従来の数理統計では、成立の怪しい仮定の下での数式展開ばかりで、多様なノイズや誤差が関与している現実のデータの処理には役に立たない。そのような現実のデータ処理を考えたデータ解析法の必要性を提唱したのであった。実際分野で Tukey が大きな成功を見ることはなかったが、これとは独立に水野坦とともに林知己夫が展開した数量化理論はその先駆となり、「囚人の仮釈放問題」や「市場調査」をはじめ、実践の場での数々の成功を見せてきた。中でも、社会調査において特に活躍するのが数量化 III 類によるパターン構造の分析である。これは、例えば 1) 日本人とハワイ日系人との意識の比較、2) 偽造回答データ検出、3) 時系列データの解析において、その有用性の認識が深まった。これについて概略を述べよう。

統計数理研究所の国際比較調査は 1971 年のハワイ日系人調査に始まったが、質問項目ごとの回答分布において、日本人と日系人との差が予想外に少なかったため、そのデータの信憑性が疑われた。日本語での調査ですら、質問の多少の表現の違いが大きな回答分布差（10～15%程度）を生じることがある。それを考慮すると翻訳を介しての比較では、その差を超えた違いを見出さねば有意味な差とはいえない。しかし、やがて時間をかけた様々な試行錯誤の後、単一項目ではなく、複数の項目群に対してその回答をクロス集計や多次元パターン解析してみると差が浮き上がってくるのが分かった。例えば日本人では質問 1 に賛成（反対）ならば質問 2 にも賛成（反対）というような一貫性がある場合、日系人では必ずしもそうではないことがあった。項目毎の単純集計表では日本人も日系人も大差ない分布であるが、項目群のパターン解析では差違が現れるのである（林知己夫・林文、1995；林 他、1998）。これは回答者集団の人々の「考え方の筋道」を具体的に浮かび上がらせるには、クロス集計やパターン解析することが必要であることを示し、またこれは「回収率と精度」の問題とは別の「データの安定性」の考えにつながる。回答分布は単純な連続量ではなく、母集団内部の各集団固有の（それぞれはそれなりの一貫性をもつ複数の）回答パターンの重畳となって、全体パターンを創出しているとすれば、回収層と未回収層との関係を考える手がかりがありそうである。つまり、未回収層も含む全体の傾向や動向を掴むのに大過はないかもしれない。（このような考えは、ポアンカレの力学系の発想を思い起こさせるかもしれない。彼は、多体問題で一般には解析的に解けぬ微分方程式の問題を、解析的に厳密に解けなくとも、その解の定性的な傾向が分かれば十分に役立つとし、その後の位相幾何学の発展に繋がった。）

また国際比較では、各国固有の事情のためサンプリング方法の差異が避けられず、海外では各種のクォータ・サンプリングを用いていることが多い。この場合、通常は標本抽出誤差をランダム・サンプリングの場合のように理論的に推定することはできない。しかし、国際比較に

における各国間の分布の差異を計るとき、林（2001, pp. 75-77）は実践的データの検討を通して、結局、項目群のパターン解析の結果はサンプリング方法の差を超えて安定していることを確認している。

一方、これまでにパネル調査の実験から、個々の人々の回答は短期間にも変動（ゆらぎ）を見せるが、集団全体（周辺分布）としては、かなり安定していることが分かっている（吉野, 1998, 2005）。その意味で、世論調査では「集団」データに関心がある。調査データが「具体的な」個々の人々の回答パターン（各個人の回答はゆらぎはあっても、その都度は1つのパターンとして一貫している）から成り立っていることを考えると、偽造データと本当の回答データは、そのパターン解析によって区別でき得るという考えに達する。実際、これはアラブ・英語バイリンガル比較調査において確認された（吉野, 1994）。初めてのアラブ人学生調査でそのデータの偽造の可能性を感じたが、アラブ語を全く分からぬ我々では、その真偽を判定することは困難で、項目毎の分布（集団平均や分散）を見て、偽造と本当の回答の差を検出することはできなかった。しかし、多次元パターン分析ではその差が浮き上がってきた。人は統計的に全くのランダム性を実現するのは難しく、偽造にも半ば無意識的にそれなりのパターンができてしまうのであろう（吉野, 1994）。

この考えを「時系列データ解析」に適用すると、人々の回答の一貫性が、時代の変遷とともに別の考え方の筋道に沿った一貫性へと遷移する場合、それを検出するのに利用できる。例えば、これは日本における昭和50年代からの近代対伝統の次元の崩壊に見られる（林, 1985）。このような概略的な解析では、大勢から外れる少数派（特に未回収層に関して）は度外視されてしまうが、我々の問題とする時間・空間的比較レベルではその程度の大勢を掴むことで必要十分であろう。

調査データの単一項目毎の分布ではなく、項目群のパターン解析や時系列・国際比較的変動の情報の把握が重要であり、その後、単純集計表に戻り、その差を再確認することが必要である。マスコミ、政治家、一般の人々への世論調査結果の説明として、パターン解析そのものを用いることは適当ではなかろうが、パターン解析等で安定した結果を確認した後に、該当部分を単純集計表等で発表することはできるのではないであろうか。

### 3.3 「標準」調査とプロファイリング

さらに、世の中全体での調査コストを考えると、標準となる時系列的標本抽出調査を確立し、それとの比較検討で個々の適時の調査データを読むことも一案である。いくつかの基本調査は共同出費で費用も時間も十分な体制で、より理想に近い標本調査として遂行し、それを「標準」（時系列変化や異なる条件の調査などの「比較基準」）として、通常は各人が個々に様々な（焦点を絞った、しかし費用も時間も限られた条件の下）調査を遂行し、その結果を標準からのバイアス等を考慮し、適正に位置づけて行けばよい。

しかし、一方で独立した様々な機関の調査による比較検討の可能性を失わないことも重要である。先の米国の大統領選挙や中間選挙では、調査コストを考えたマスコミ各社が1つの調査会社に任せ、結局、予測が失敗し、さらにはその失敗の原因究明もできない事態のまま、その会社もつぶれてしまった。この事例が示すように、各社別個の調査結果を見て、それらの内に何らかの一貫した傾向（不変量）を確認することは必要である。

特に重要なのは、常に具体的な「生きた人間の考え方の筋道」を浮かび上がらせること、いわば社会調査による集団のプロファイリングである。例えば、「無党派層」の行動様式を読むことや、選挙結果の予測調査に「未回収層」になりやすい人々の性・年齢・学歴・職業などの属性（これはセンサス・データとの比較で分かる）、行動様式、性格等々を常時、何らかの方法で掴んでおくことである。ケース・スタディ的な調査の積み重ねも重要となろう。これによ

り、たとえ全体の一部であっても確固たる構造を把握できていれば、全体を再構成（回収層データから、非回収層をも含む全体データを推測）するにあたって本質的な間違いは少なくなるであろう。（考古学的骨片から恐竜の全体像を復元し、認知科学の計算的視覚理論で木の間に見える象の一部から象の全体を「見る」のと同じ発想で、これは、ランダムサンプルを基礎とする思想とはかなり異なる。たとえわずか一部の「偏った」情報でも、全体を推定するのに十分な情報となることがある。これが成立するには、対象のダイナミズムが理解できていて、その一部の情報が全体のどこに位置づけられるかが明確になっていなければならない。）いわば、従来の大標本調査とプロファイリング（ケース・スタディ的調査）の相補的データ収集が必要とされる。

### 3.4 「世論」のダイナミズムを読む

経済では均衡理論というのがある。経済状況は、常にある均衡点（安定点）にあるか、またはそれに向かって変化しているというのである。これを模して、世論も変動はするが、各時点毎にはそれぞれある特定の均衡点（回答分布）にあるかそれに向かって変化しているとみることが可能である（複雑系の理論）（吉野, 2002a; 吉野編, 2007）。この考えも、回収層と非回収層の回答パターンの関係を考えるのに役立ち得る。現実には、その均衡点を計算する問題は容易ではないだろうが、不可能とも思えない。課題は、実践的方法によりどこまで有効な情報が把握できるかである（Yoshino, 1998, 2002a）。現実には、モニターによるトレンド調査やパネル調査など、この方面の努力が続けられている。1回毎の調査は母集団全体から見るとかなりのバイアスがあろうが、時系列的変化の情報、従来の標本抽出理論を超えた情報を与えている。

## 4. 新しい時代の「世論調査」のために

これまで述べたことは、個々の結論だけを見れば必ずしも目新しいものではなかろう。しかし、戦後の先輩諸氏が尽力したように、それらを実践を通して理論づけ、従来の（静的）標本抽出理論に代わり、日本全体の調査関係者の合意の下で新しい時代の世論調査の基盤（「定義」の見直し）を作ることが必要に思える。「壺から色つきビー玉をランダムにとり出し、全体を推定」する従来の標本抽出理論ではなく、社会のダイナミズムを読む世論調査（集団力学）への昇華である。

本論文の前半では、少し厳しく「世論調査」の現状を批判した。しかし、これは先人たちの歴史を尊重した、統計数理研究所の「役割演技」と考えて受け入れていただきたい。統計数理研究所すら科学的調査を大義とする態度を崩してしまう時、それは日本の世論調査とそれに従事する専門家者たちの凋落を意味する。「悪貨は良貨を駆逐」し、標本の代表性の不明なままのコンピュータ・ネットによる「アンケート」が日本を覆うこととなり、これまで現場で真摯に調査に従事してきた官民学の調査機関をつぶすことになる。我々は、最低限、譲ることのできない条件を守らねばならない。

その一方で、現場の世論調査遂行の本音と建前を十分に承知し、実践可能な方法や条件の具体的提案をする任務がある。「方法」については不十分ながらもすでに述べたが、「条件」については調査費用に触れて結びに代えよう。

この2、3年で官公庁では社会的アカウントビリティのためと称し、調査にも入札制度を適用するようになってきた。慎重な条件提示で時間をかけた入札が可能であれば別かもしれないが、それを心得ていない、金額の多寡のみにしか注意を払わない依頼者の場合は、大きな不都合が生じる。見かけ上は同様なデータでも、十分に収集方法やクリーニングに配慮したデータとそうでないものとの質の差は大きい。我々にはそれを見極める能力はあるが、しばしば、そうでない人々の間でデータの数値は一人歩きしていく。これは、実際に官公庁でも起きてい

るようである。良心的な調査会社がほとんど利益なしで調査を遂行しているか、民主主義を守る大義には無関心の会社がコスト内で低品質の調査を遂行しているのが現状に近い。

かつて筆者は、故・林知己夫とともに大きな科学研究費プロジェクトを提案したことがあった。その際、審査委員会で調査研究を知らぬ委員から「なぜ社会調査にそれほどの費用がかかるのか？」と疑義が出たそうであり、結局、その計画は実らなかった。しかし、その後、「ニューヨークの9月11日テロ」が起き、中東を始め、地域調査の重要性が急に叫ばれ始めた。しかし、事が起きてからでは、我々のような調査はできないのである。調査研究者は常に「治にいて乱を忘れず、乱にいて治を忘れず」でなければならない。戦闘機1機にはるかに及ばぬ費用で世界の状況を相互理解し、各国の民主主義を守っていくのは、費用対効果から考えても格段に優っているのである。

また、世論調査の方法論では日本は世界中で最も民主主義的であるかもしれないが、他方でデータ公開の点では先進国では一番遅れているという点が看過できない。これに関係して、社会調査士資格認定機構が2004年10月（直井，2003）に発足し、2008年度の法人化に向けて動いている。また、学会の統計関連の分野では「政府統計・社会統計情報基盤整備分科会」（廣松毅委員長）、社会学の分野では「社会調査分科会」（盛山和夫委員長）が設けられ、データ公開の促進を含め、調査統計の社会貢献について具体的検討が進められている。

それらが確実な成果へ結びつくには時間もかかるが、これまで「科学的世論調査」を守って来た各位の御協力で、世論調査あるいは統計的標本抽出法に基づく社会調査の意義が広く一般社会に理解されていき、これまで調査拒否をしてきた人々が、意見を表明することが民主主義の社会参画にとって重要であり、また権利でもあると積極的に考え、結果として調査データの質も向上することが望まれる。そして、日本でも諸外国同様に本格的なデータ・アーカイヴが設立され、統計調査データが広範に官民学を含む、一般の方々の利用に供され、世界の安定した平和的発展と経済繁栄の基礎情報として活用されることを期待する。

## 謝 辞

本論文は、輿論科学協会の発刊している「市場調査」, No. 259, pp. 4-13 (2004) に発表した「科学的世論調査法の価値」を、同協会の許可を得て、その後の社会調査環境の状況と本特集号の趣旨を勘案して加筆修正したものである。同協会のご協力に感謝いたします。また、本論文の原稿について、匿名の審査員から貴重なコメントをいただき、感謝いたします。その基本的な部分は統計的推測理論やRDDやインターネット調査等、各種調査法の多様な可能性についてのご示唆であり、今後とも世論調査研究者が検討すべきことでありました。我々は、一方でそのような理論を発展させる努力を続けていく課題と、他方で、それを必ずしも統計の専門家だけではない一般社会がどのように受け入れていくべきかを考えるという課題の双方を実践的に捉える必要があると思われます。本論文では、最近の調査環境の悪化で難題を抱えた分野の今後の発展を期待して、多少とも、主観的な檄文をも交えたことはご容赦願います。

## 参 考 文 献

- アンガー, J. M. (2001). 『占領下日本の表記改革』奥村陸世訳, 三元社.  
 エルドリッジ, R. D. (2003). 『奄美返還と日米関係』南方新社.  
 林知己夫 (1977). 『世論調査の現状と課題』「第2章世論調査の発展と現状」輿論科学協会編, 至誠堂.  
 林知己夫 (1984). 『調査の科学』講談社ブルーバックス.  
 林知己夫 (1993b). 『行動計量学序説』朝倉書店.  
 林知己夫, 林文 (1995). 「国民性の国際比較」『統計数理』43, 1, 27-80.  
 林知己夫 (2000). 「これからの国民性研究—人間研究の立場と地域研究・国際比較研究から計量的文明論の構築へ—」『統計数理』48, 1, 33-66.

- 林知己夫 (2001). 『データの科学』朝倉書店.
- 林知己夫編 (2002). 『社会調査ハンドブック』朝倉書店.
- 林知己夫, 鮑戸弘 (1976). 『多次元尺度解析法』サイエンス社.
- 林知己夫, 水野坦, 青山博次郎 (1952). 『数量化と予測』東京丸善株式会社.
- 林文・田中愛治 (1996). 「面接調査と電話調査の比較の一断面—読売新聞社世論調査室の比較実験調査から—」『行動計量学』23, 1, 10-19.
- 本田則恵, 本川明 (2005). 「インターネット調査は社会調査に利用できるか」『労働政策研究報告書』労働政策研究・研修機構.
- Inkeles, A. (1996). *National character*, Transaction Publishers, NJ. 吉野諒三訳 (2003). 『国民性論』出光書店.
- 小林和夫 (2003). 「世論調査公表の自由」世論調査協会大会発表.
- 高倉節子 (2003). 「林先生の足跡を訪ねて」『行動計量学』30, 2, 193-197.
- 高橋正樹 (2004). 「科学史と科学者—林知己夫氏公開インタビュー—」『行動計量学』31, 2, 107-124.
- Tukey, J. (1977). *Exploratory Data Analysis*, Addison-Wesley MA.
- Little, R. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data 2<sup>nd</sup> ed.* John Wiley & Sons.
- 森本栄一 (2005). 「戦後日本の統計学の発達—数量化理論の形成から定着へ—」『行動計量学』32, 1, 45-68.
- 水野坦 (2003). 「統計数理とは何か」(昭和26年より朝倉書店から刊行された「統計数理講座」の葉より再録), 『行動計量学』30, 2, 191-192
- 村井実[訳] (1979). 『アメリカ教育使節団報告書』講談社学術文庫 253.
- 直井優 (2003). 「社会調査士資格制度の成立」『日本世論調査協会報』92, 58-60.
- Narens, L. (1985). *Abstract Measurement Theory*, MIT Press.
- ラオ, C. R. (藤枝康祝, 柳井晴夫, 田栗正章訳) (1993). 『統計学とは何か—偶然を生かす—』丸善株式会社.
- 鈴木督久 (2006). 「エリアサンプリング調査の再検討」『日本行動計量学会 34 回大会発表論文抄録集』286-289.
- 山本武利 (2001). 『日本兵捕虜は何をしやべったか』文春新書 214.
- 読み書き能力調査委員会編 (1951). 『日本人の読み書き能力調査』東京大学出版会.
- 吉野諒三 (1989). 「公理的測定理論の歴史と展望」『心理学評論』132 (2), 119-135.
- Yoshino, R. (1998). A social quantum theory for the analysis of public opinion survey data, *Behaviormetrika*, 25, 2, 111-132.
- 吉野諒三 (1994). 「国民性意識の国際比較調査研究—統計数理研究所による社会調査研究の時間・空間的拡大—」『統計数理』42, 259-276.
- 吉野諒三 (2001). 『心を測る—個と集団の意識の科学—』朝倉書店.
- 吉野諒三 (2002a). 階層的2進分解モデル (HiBiD) の構成—個と集団の認知の力動的解析の為の「社会的量子論」『認知科学』9, 2, 213-229.
- 吉野諒三 (2002b). 「平成13年度 世論調査に関する調査研究—世論調査のサンプリング方法について—研究報告書」内閣府大臣官房政府広報室.
- 吉野諒三 (2003). 「日本における国民性研究の系譜」『国民性論』Inkeles, A., 1996 (吉野訳, 2003) の付章. 出光書店.
- 吉野諒三 (2005a). 「東アジア価値観調査—文化多様体解析 (CULMAN) に基づく計量文明論の構築へ向けて—」『行動計量学』32, 1, 133-146.
- 吉野諒三 (2005b). 「富国信頼の時代へ—東アジア価値観国際比較調査における信頼感の統計科学的解析—」『行動計量学』32, 1, 147-160.
- 吉野諒三 (2006). 「今世論調査が直面する壁—「歴史」と「理論」と「実践」—」日本世論調査協会報『よろん』97, 31-36.
- 吉野諒三 (2007). 「海外の標本抽出面接調査の方法, いんぷおるむ (第53回)」『新情報』95, 7-12.
- 吉野諒三 (編) (2006). 『環太平洋 (アジア・太平洋) 価値観国際比較調査における信頼感の統計科学的解析—中国2005年調査—』統計数理研究所.
- 吉野諒三 (編) (2007). 『数理心理学』心理学の世界 専門編 16, 培風館.
- 吉野諒三 (編) (2007). 『東アジア国民性比較データの科学』勉誠出版.
- Yoshino, R. and Hayashi, C. (2002). An overview of cultural link analysis of national character, *Behaviormetrika*, 29, 2, 125-142.