

林知己夫先生と多次元データ解析

—数量化法、データ解析、そして分類からデータの科学へ—

統計数理研究所調査実験解析研究系教授

大隅 昇

林知己夫先生がご逝去された。まさに「巨星墜つる」のひと言に尽きる。統計学にかぎらず、社会調査、データ解析の諸分野での今のこの混沌とし、また危機的な状況にある今、我われにとっての羅針盤であり、なくてはならない先導者であった先生を失ったことは無念でならない。

実は先生は、私の恩師であった故宇野利雄先生の追悼文の初めに「学問上でも人格の点でも私の最も尊敬する先生です」と記されている。宇野先生は、数値解析・数値計算論の先駆者であり、また優れた教育者でもあったが、正に林先生に対して、これと同じ言葉を捧げたい。

言うまでもなく、先生の膨大な研究業績のすべてをここに述べることはできない。先生との長いおつき合いの中での様々な思い出も同様で、これも書きだせば際限がない。ここでは、私からみた林先生を、またおそらくは、日頃ほとんど知られることのなかった事々の断片を、追悼の意味を込めて書き記したい。

◆好奇心旺盛な先生から学んだこと◆

先生は、博識多才、非凡にして希有な存在の方であった。また、音楽、スポーツ、書、茶道…と多彩なご趣味の方でもあった。そして何より、大変に好奇心旺盛で、新しいことが何でもお好きであった。私事で恐縮だが、私が20数年前に当時注目され始めたファジイ理論を分類法に応用した研

究（ファジイ・クラスタリング）で学位取得を考え、林先生に審査委員をお願いに伺ったときのことである。先生は開口一番、「これは面白い、自分も目を付けて調べていたところである、是非これを進めてみよ」とのお言葉をいただいた。当時、ファジイ理論とは一体どんなものかがほとんど知られておらず、これをすでにご存知であったことにひどく驚いた。これに限らず、何か新しい話題が登場すると、研究室に電話があり、「おい、○○とかいう何やら分からぬ研究があるようだが、君は知っているか」「これこれの情報や資料があれば届けてくれ」ということが頻繁にあった。今思い出すだけでも、フラクタル、遺伝的アルゴリズム、ニューラルネットワーク、ラフ集合、複雑系、サポートベクターマシン（SVM）、そしてデータマイニング（DM）に知識発見（KD）等々、世情で話題になっていることの大半は、目を通され、また必ず何らかの意見をお持ちであった。先生は人を使いこなす達人で、おそらくは、あちこちに人脉ネットワークを張り巡らし、常に最先端の情報収集に専念されていたのではなかろうか。

私はこれを、研究者たるもののが心地、とくに統計学やデータ解析の研究をめざす者としての目配りのありようを、先生自ら示されたメッセージととらえてきた。確かに、先生の言葉にそって物事を眺めると、情報も別の形として見えてくるもの

で、研究者として大いにありがたいことであった。今にして思うと、これは先生独特のご指導の方法であったようだ。

◆多次元データ解析と多変量解析は異なるもの◆

先生は、日頃、物事の本質はその言葉や書きようで端的に表されるものと強く考えられていた。論文の内容はもとよりであるが、専門用語の使い方にも厳しく、例えば、統計科学は好まず、統計学であるとおっしゃる。同じように、多変量解析ではなく、多次元データ解析であると言われる。我われ凡才には、とりあえずはどちらでも…と考えるのであるが、先生はそれではすまない。ある雑誌の多変量解析に関わる特集記事で、私の記事の通読をお願いしたが、うっかり多変量解析とした箇所はすべてバツ印が付けられており、「これだからいかん」と注まで付いていたのには驚いた。同時に、気に入ったところには「然り」とあり、ご自分の記事に考え方方が「合っておるのでよし」となる。これらの使い分けが絶妙で、書き手を思わず納得させるほど、その指摘や示唆が的確なのである。同時に、些か誤字や脱字があっても意に介せず、分析や論旨の本質的な部分で筋が通っておればよしとの精神であったように思う。これも、先生流のご指導であったのだろう。ともあれ、先生にとっては、多変量解析とはガウス分布（正規分布）に基づいたらずに複雑な算術操作にすぎず、あまりに制約が多く実用には耐えないもの、せいぜい論文作成の装置にすぎないと考えておられた。そして「木を見て森を見ず」があまりに多いと、たえず嘆いておられたことを思い出す。

◆現象解明のための数量化◆

林の数量化理論と一般に言われているが、意外なことに先生は、その多数のご著書の中で、単に「数量化」と書かれている場合が多い。また、数量化のⅠ類、Ⅱ類、Ⅲ類、Ⅳ類、さらにはⅤ類、

VI類といった命名は飽戸弘先生によるもので、林先生ご自身はそれぞれの考え方を映すような別の呼び名を与えていた。しかし、世の中への普及のためには、飽戸先生の命名が大いに役立ったとも折に触れ述べられている。こうした細かい気配りには敬服するばかりである。

数量化の原点は「数（数値）はあらかじめ存在するものではなく、我われの目的に応じてふさわしく与えるもの」という思想に依拠している。これは、そもそも生の測定データ（数値とは限らない）の示す意味表現と、分析に用いるために必要とする数値とは峻別して考えるべき、との見方である。俗に、質的あるいは定性的データに数量を与えるものとの言い方があるが、それと同時に、元来は非線形の世界の物事が多いのであるから、それをなるべく扱いやすい線形にすること、合わせて実験の計画を工夫し、その現象解明に適したデータの取得法と解析法を通じて問題を解明する筋道を明らかにするという立場である。こうしたことは、精緻なモデリングを前提とし、それにデータを合わせるような分析姿勢に対するアンチテーゼであって、終生、大方の統計家とは大変に異なる立場をとられた。また、このことが本来の統計学の分野からは、先生のこうした精神は正しく理解されなかつたのではなかろうか。同時にこのことが、晩年の「データの科学」への強い主張となつて表れたようにも思う。

◆ベンゼクリ教授との出会い、そして対応分析◆

フランスに、ベンゼクリ（J.-P. Benzécri）という研究者がいる。彼は1960年代初期に、レンヌ大学で、初めてAFC（Analyse des factorielles correspondances、いわゆる correspondence analysis：対応分析）という方法を言語学のデータに適用した事例を発表する。そしてこれこそがフランスにおけるデータ解析（analyse des données）の原点であるとし、また、従来の数理統計的アプ

ローチを強く批判し、パリ第6大学における彼の独特的講義は広く知られるところとなった。また、講義だけではなく、多数の著書の中で（それも専らフランス語による）、フランス流のデータ解析のありよう、思想を主張するのである。AFCは実は数量化法Ⅲ類と同じではないかとの指摘があるものの、確かめようがない。林先生は、果たしてベンゼクリは実在するのか、ブルバキと同様に研究者集団ではないのかとまでいわれ、探偵もどきの探索までされた。しかしそれも、1979年に林先生とベンゼクリ教授との初めての出会いが実現し、ひとつの節目を迎える。前年の1978年に、日本学術振興会（JSPS）の援助で、松下嘉米男先生（当時、統計数理研究所部長）を日本側代表とする数名の研究者（西平重喜先生、杉山明子先生ほか、筆書も含む）とパリ第6大学統計学研究所の（故）デュゲ教授との間で、日仏科学協力セミナーが開催されたのである。そしてこの席で、一同、初めてベンゼクリ教授に会う機会を得る。この情報は直ちに林先生に伝えられ、次の年の初顔合わせとなるのである。ベンゼクリ教授は、ある著書の中で、「まさか遠い東洋の国に、自分よりもはるかに早く、同じようなことを考える研究者がいたとは驚きであり、また敬服に値する」と述べている。こうして、フランス圏はもとより、広く欧州圏で数量化法Ⅲ類が対応分析と同等の方法であることが知られることになった。しかし、林先生は、数式や定式化のうえでの同等性はあっても、その根底にある思想や哲学はまったく異なると主張される。事実、フランスでは対応分析は定性データの主成分分析といった趣で議論され、これを巡る数理展開は理屈好きのフランス人らしく、実に詳細かつ緻密に調べられている。しかし、林流のように、質的データ全体に対して俯瞰的であり、また首尾一貫した“数量化”という独創的な発想とは次元を異にするのである。ここでも、やはり林先生独自の定性情報の分析に対する先見

的な思想は優れて希有であり、またこの考え方は広く海外に知られるようになった。

◆日仏研究交流—データ解析を巡る国際交流—◆

周知のように、「データ解析」(data analysis)、とくにEDA（探索的データ解析）は、テュウキー博士（J. Tukey）が1960年代の初期、つまりAFCの登場とほぼ同時期に提唱した考え方である。林先生は早くからこれに注目し、先生の考えるデータ解析に相通じるものがあるが、いかにも限られたデータしか頭にないようだと言っておられた。先生の考えるデータ解析とは実務・応用に密接に繋がった現象解析解明の手段としての活きた方法論であり、また数量化そのものであらねばならなかった。

ともあれ数量化のこと（とくに、数量化法Ⅲ類）が国際的に広まるとともに、また、ベンゼクリ教授との接点もできたことで、フランス圏の研究者グループから林先生を招きたいとの声が高まり、ついに1979年にそれが実現する。当時、INRIA (Institut National de la Recherche en Informatique et en Automatique) の研究者であったディディイ教授 (Diday) 他、ベンゼクリ派のフランス流データ解析の精鋭部隊が、ベルサイユで隔年に開催してきた「データ解析と情報学」(Data Analysis and Informatics) の国際研究集会への招待講演者として林先生を招くことになった。このときの先生の講演は今でも鮮やかに覚えている。無礼を省みず申せば、先生の英語のトークは決して上手ではない。しかも、例によって先生流の哲学的な含意の多いお話である。先生の話が始まると会場はざわめき、しかし発表後は拍手で溢れ、遠い東の国のカリスマの登場を歓迎する声で満ちた。ベンゼクリの弟子達が次々と握手を求め、先生の数量化の思想はフランスのデータ解析に相通じるものであるとの評価がなされた瞬間である。

先生は、その後、ベンゼクリ教授を日本に招聘

したいと何度も働きかけたが、彼が航空機による旅行は一切しないとの理由で、彼の弟子筋にあたるルウ教授 (M. Roux) を代わりとして招聘することになる。これを契機として、フランス圏の多数の研究者との交流が始まり、ルバール (Lebart)、ジャンブ (Jambu)、ディディイ (Diday)、ナカシュ (Nakache)、エスクフィエ (Escoufier) 等々、現在のフランスのデータ解析を先導する人達であり、またベンゼクリ派の門下生を含む多数の研究者との長い交流が続くのである。とくに先生は、ルバール教授とは、国際比較調査や国際会議での共同研究発表など、深い絆を築かれ、また彼をはじめ多数の研究者を日本に招くなど努められた。我われもこの恩恵にあずかり、いまなお2国間の深い研究交流が続いている。またデータ解析の分野だけでなく、意識の国際比較方法論の研究の分野でも、広く海外の研究者との交流、共同研究を積極的に進めてこられた。

先生は、常にこうした人脈を大切な資産とし、様々な活動を展開された。とくに日仏科学協力セミナー開催を、我われとともに組織的に進め、日本学術振興会 (JSPS) とフランス国立科学研究中心 (CNRS) との支援で、1987年（東京）と1992年（モンペリエ）の2度にわたり実現することとなる。先生はこうした学界活動にも常に精力的に取り組まれ、また我われをたえず励まし、刺激することを忘れなかった。

◆分類学研究からデータの科学へ◆

先生の国際的な研究交流は、フランス圏の研究者から次第に、英国、イタリア、ドイツ、ポルトガル、そして東欧圏諸国へと拡がるのであるが、そのきっかけのひとつは分類学研究にある。先生は日頃から「“分類”はあらゆる科学における基本的な思考操作である」とされ、実際に数量化を考えるうえでも分類操作の重要性を述べられてきた。こうした研究分野に関わる研究者の集まりが、あ

ちこちの国に現れ始め、とくに北米分類学会と英國分類学会の発足に伴い、フランス・ベルギー圏、ドイツ、イタリア、スロベニア、ポーランド等に次々と分類学会が設立された。林先生は生来的好奇心と負けず嫌いからであろうか、我われも分類学会を作ろうと言いだし、まずは研究会形式で発足させた。これが1983年のことである。その後、1991年に学会組織とし、現在に至るのである。その後の国際分類学会連合 (IFCS : International Federation of Classification Societies) への参加、そしてIFCS国際会議（第5回大会、1996年）を日本（神戸）で開催することへと繋がる。ここで先生は、初めて「データの科学」についての基調講演をされ、その理念と重要性について述べられた。それは、先生流の探索的データ解析のひとつの発展形であり、何よりもデータこそが現象説明の礎であり、それをいかに取得し、解析するかを独自に考えることこそが原点であるとするものである。

このように、欧州圏の多くの著名な研究者から、その業績を高く評価され、英國王立統計協会名誉会員、国際統計協会 (ISI : International Statistical Institute) 正会員、国際分類学会連合会長などを歴任された。

晩年の十数年は、我われとともに「科学方法論としてのデータの科学」の重要性を、先生独特の語り口で主張されていた。話の順が逆となるが、実はこれの語源は、第2回日仏科学協力セミナー (1992) の論文集発刊の際、フランス側の代表であるエスクフィエ教授と、その表題や序文を議論する中で生まれた言葉「data science」にある（論文集は “Data Science and its Applications” となった）。先生は、これに日本語の「データの科学」をあてられ、終生この言葉でその思想を語られた。また、この第2回集会の論文集を発刊されるとき、数量化に関わる用語を初めて「数量化法 (quantification method)」で統一したいと

提奏され、しかも以後の英文の表記はこれで定められるようになった。我われがデータ科学と言ふ、私のグループがデータ・サイエンスと言うと、それは「思想が違う」とお叱りを受けたことも、今はや懐かしい思い出となってしまった。

データの科学の根幹は、至極まつとうな理念である。それをあえて主張せねばならなかつたことこそ、今の統計学や統計的データ解析の世界の混沌、さらには調査環境の悪化と、先の見えない危機的な閉塞感への警鐘であり、これをいかにブレイクスルーするかを模索する道筋を示されたかったに違ひない。

こうした先生の永年にわたる偉大な業績に対し、昭和56年に紫綬褒賞、平成元年に勲二等瑞宝章が授与され、また没後、正四位を叙された。

このようなことで、先生と私との接点から眺めた思い出や心情を述べることとなつた。またここに記したことは、先生とのおよそ30年にわたる長いおつき合いの中での、わずかな断片でしかない。

社会調査ハンドブック

林 知己夫 編集

A5判 776頁 本体25000円(12150-4)

マーケティング、選挙、世論、インターネット。調査のニーズはますます高まっている。本書は理論・方法から各種の具体例まで、社会調査のすべてを集大成。調査の「現場」に豊富な経験をもつ執筆陣が、ユーザーに向けて実用的に解説。
[内容]社会調査の目的／対象の決定／データ獲得法／各種の調査法とそれを行う方法／調査のデザイン／質問・質問票の作り方／調査の実施／データの質の検討／分析に入る前に／分析／データの共同利用／報告書／実際の調査例／付録：調査のための基礎データの獲得法／他

朝倉書店

〒162-8707 東京都新宿区新小川町6-29
電話 営業部(03)3260-7631 FAX(03)3260-0180
<http://www.asakura.co.jp>

*本体価格は消費税別です。
(ISBN)は4-254-を省略
*ホームページで「書籍注文」ができます

再び私事となるが、フランスのロワール河・城巡りの小ツアーで同じ部屋に泊まつたこと、モロッコで猛暑の中、先生ご夫妻とエクスカーションと共にし、また絨毯の値定めをしたこと、ローマの暑い日盛りの中を、一緒に探索して歩いたこと、外国のレストランで何度もご相伴にあづかったことと、あれこれが走馬燈のように巡る中、いまにもベレー帽の先生が現れて、しっかり頑張れ、頼むぞ、と励ましてくださるのではという気持ちで胸がいっぱいになる。

先生は行年84歳であった。もっと生きていただき、今の統計学やその周辺にみられる、この危機的状況から脱却し、望ましい方向へ進むための優れたお知恵を頂き、道標となつていただきたかった。かえすがえす無念でならない。

終わりに、今後のデータの科学の理念の普及発展をめざすエヴァンジェリストとなるべく、また新たな調査法の研究に努めることに誇りを持って、泉下にある先生のご冥福を衷心よりお祈り申し上げる。合掌

シリーズ〈データの科学〉

林 知己夫編集

1. データの科学

林 知己夫著 A5判 144頁 本体2600円(12724-3)
[内容]科学方法論／データをとること／データを分析すること／質の検討・簡単な統計量分析からデータの構造発見

2. 調査の実際

林 文・山岡和枝著 A5判 224頁 本体3500円(12725-1)
[内容]データの獲得／どう調査するか／質問票／精度。データから情報を読みとる／データの特性に基づいた解析／他

3. 複雑現象を量る

羽生和紀・岸野洋久著 A5判 176頁 本体2800円(12727-8)
[内容]紙リサイクル社会／背景／文献調査／世界のリサイクル／業界紙に見る／関係者／資源回収と消費／他

4. 心を測る

吉野諒三著 A5判 168頁 本体2800円(12728-6)
[内容]国際比較調査／標本抽出／調査の実施／調査票の翻訳・再翻訳／分析の実際(方法、社会調査の危機、他)／他

5. 文化を計る

村上征勝著 A5判 144頁 本体2800円(12729-4)
[内容]文化を計る／現象解析のためのデータ／現象理解のためのデータ分析法／文を計る／美を計る／古代を計る／他

—— 総 刊 ——

6. データの科学とデータマイニング 大隅 界・吉村 実著

特集

「小売業の顧客囲い込み戦略の新潮流」

特集の主旨	橋本谷 智子	2
ウォルマートのグローバル戦略：日本市場参入への意味合い	小田部 正明	3
小売業の顧客囲い込み戦略とリサーチの活用	清水 正博	14
マーケティングの新手法・ゲーポンサービス	小山 義彦	21
小売業との協働による顧客囲い込み事例の報告	真田 久仁彦	31

“林 知己夫先生が残されたもの”	飽戸 弘	38
林 知己夫先生と多次元データ解析 －数量化法、データ解析、そして分類からデータの科学へ－	大隅 昇	41
故 林 知己夫先生の業績を讃える	小林 和夫	46
中国・上海事情- 9	上田 滋	49
米国・ニューヨーク事情- 3	内田 俊一	51
ジャーナルレビュー（国内・海外）	平松 貞實／小林 和夫	54
企業訪問レポート－日本統計調査(株)－	豊田 みゆき／武田 修	55
投稿の呼びかけ・投稿規程	出版委員会	50
編集後記	武田 修／山下 育夫／橋本谷 智子	58