

《研究ノート》

プッシュボタン式統計調査の効用と限界 (2)

——神と悪魔はコンマ以下に宿る？——

佐 藤 郁 哉

承前——Excel（や統計解析ソフトウェア）を使うか，Excel に使われるか？

- I 測定精度過信症の症例と病像（診察①：What）
 - II 測定精度過信症と「やっている感」の受容（診察②：What）
 - III 不可解なグラフと測定精度過信症の背景（診断：Why）
 - IV 「こたつ論文」からの脱却（処方箋：How to）
- 【補論】 データ縮約の「べし・べからず集」（Do's and Don'ts）

正しい問い——これ自体しばしば曖昧なのだが——に対するおおよその答えは，間違った問いに対する厳密な答えよりもはるかに良い。というのも，後者は幾らでも「不必要なほどに」精確なものにすることができるからである（Tukey 1962: 13-14. 強調は原文）。

承前——Excel（や統計解析ソフトウェア）を使うか，Excel に使われるか？

1. 便利なソフトウェアの罠

前稿では筆者が20年以上前に遭遇した，全体が鮮やかな青色で塗り尽くされた「100%の円グラフ」のエピソードを話のマクラにして，円グラフ，折れ線グラフ，帯グラフという3種類の図解表現の使用をめぐるある種の誤解とその背景について検討を加えた。その検討の結果として浮かび上がってきたのは，論文等についてその書き手の都合だけが優先される一方で読み手の立場や視点があまり考慮されていない，という状況である。つまり，そのようなケースでは，グラフというものが本来持っているはずの，円滑で効果的なコミュニケーションを促進するための用具という点に関する認識が欠如しがちになっているのである。

また前稿では，Excel や Google Forms などをはじめとする「プッシュボタン式」のプログラムやサービスは，基本的な集計や統計解析をおこなう際には便利この上ないツールではあるが，ともすれば，不適切で不可解な図解の作成に結びついていく可能性がある，という点についても指摘した。実際，例えば Excel の初期設定で各種のグラフを作成すると，数値データの構造や意味がむしろ読み取りにくくなる夾雑物が自動的に盛り込まれてしまう場合が多い。

つまり、我々は便利な表計算ソフト等を自在に使いこなしているように見えて、実は、その種のコンピュータ・プログラムによっていわば「マインド・コントロール」されている場合が少なくないのである。前稿でコール・ナフリック (Cole N. Knaflic) による6つの教訓を参考にして指摘したように、より明快な図解表現を目指す際には、Excelをはじめとするソフトウェアの使用法について考え直していく必要がある。その点をひと言で言い表せば次のようになるだろう——「Excel を使うか、Excel に使われるか?」。

2. 対症療法から根治療法へ

当然ではあるが、ソフトウェアとの「つき合い方」をめぐる問題は、Excelに限らず例えば各種の統計解析サービスやデータ・マイニングないしテキスト・マイニング関連のソフトウェアなどについても指摘できる。実際、特に深く考えることもなく、その種のプログラムに組み込まれている既定の手続きやフォーマットに従っている限りは、「読者フレンドリー」を心がける上で本来必要になるはずの思考過程が端折られてしまうことが稀ではない。いかにも専門的で科学的かつファッショナブルな用語や小数点以下の数値が並ぶ係数、あるいは色とりどりの印象的な図形を散りばめた図解表現は、ともすれば不徹底な分析を覆い隠す一種の「目くらまし」になってしまう場合が少なくないのである。

もっとも、その目くらまし的な図解がしばしば数値データの解析を中心とする「科学的で客観的」な定量的調査の分析結果として通用してきた、というのもまた紛れもない事実である。だからこそ、例えば『下流社会』(三浦2005)はベストセラーとなり、また「市場調査を土台とするしっかりしたデータにもとづいて、低収入層の生活スタイルを分析」(野口2005)した書籍として評価されることもあったと考えられる。

本稿ではまず、数値データの図解というよりは数値データそれ自体の提示法、とりわけ小数点以下の数値の「精度」を典型的な事例ないし「症例」としてとりあげて検討していくことにする。ここで小数点以下の数値の精度を事例として取りあげるのは、これが数値データの科学性や客観性を盲信してしまう「数値信仰」を明確かつ象徴的に示す事例として考えられるからに他ならない。

前稿では、日常生活でも目にすることが多い3種類のグラフ例をとりあげて、データを視覚化する際の典型的な誤用例を取りあげた上で、それぞれに関する改善策を提案した。医療に喩えて言えば、それらの改善提案は当座の症状を緩和させるための「対症療法」だと言える。

それに対して本稿では、不必要な小数点以下の数値という事例(症例)を切り口にして、次にあげる2つの問題について検討していく。

- ① (「物理学への羨望」に象徴される) 厳密社会科学の幻想
- ② データの「縮約」(data reduction) に関する誤解

これらはいずれも、我々の思考の中に根強く存在する数値信仰という「宿痼」と密接に関連するものである。その意味では、この2つの問題について検討していくことは、不適切な図解表現や数値表現の背景にある本質的な原因(病因)について明らかにし、またより根本的な改善策＝根治療法を模索していくことに他ならない。

I 測定精度過信症の症例と病像 (診察①: What)¹⁾

アンケート調査などで単に「(科学的な調査を) やっている感」を演出するためであれば、各種のグラフなどの図解表現だけでなく小数点以下の桁を含む数値も効果的であろう。実際、「コンマ以下」の桁数が多ければ多いほど、いかにも精度の高い測定器具がはじき出した客観的な数値のような印象を与えることができる。その種の数値は単純な測定値の場合もあれば、測定値を集計した結果を示す値(平均、分散、比率、各種の指数係数値等)であることも多い。また、小数点以下を含む数値がグラフと組み合わせて使用される場合も少なくない。

そのいずれの場合にせよ、小数点以下の数値が含まれることによって、いかにも精度の高い、客観的で科学的な測定がなされているという外見を取りつくろうことができる。また、人間の主観を排除して厳密な数学的手続きを経て算出された値であるという印象を醸し出すこともできる。しかし、物理現象などの場合とは異なり、社会現象に関する測定については必ずしもその精度が保証されているとは限らない。それどころか、その測定の精度はおろか測定法自体の詳細が明確には示されていない場合さえ少なくないのである。

1. 大学ランキング

(1) 症例1:『ゴーマンレポート』——小数点以下2桁の数値で違ってしまう順位

測定の基準や方法を公開していないにもかかわらず、小数点以下の数値を用いて(濫用して)社会的に大きな影響を与えかねない重要な主張をしていたものとして最も有名な(悪名高い)例の1つに『ゴーマンレポート (Gourman Report)』がある。これは、主に米国の高等教育機関のランキングを掲載した冊子であり、1967年の第1版以来1997年の第10版まで30年近くにわたって刊行されていた。著者は、大学教員としての勤務経験もあるジャック・ゴーマンである。最初は自費出版で刊行されていたが、後にテスト会

1 本節および次節の一部は、佐藤(2021: 293-300)を元になっている。

社であるプリンスン・レビュー社から刊行され、大手出版社であるランダムハウスが流通を請け負っていた。また、邦訳が刊行されたこともある。

このランキングについて問題視されていたのは、順位付けをおこなった際に使用した方法について「我々がある基準で選んだ人々からもたらされた膨大な史料に基づいている」(ゴーマン 1998: xii) などとするだけで、著者のゴーマンはその詳細については公開していなかったという点である。

例えば、1997年版のゴーマン・レポート(ゴーマン 1998)では、〈評価に際しては各大学の教職員から提供された内部資料以外に外部資料も参考にした〉と述べ、また18の評価基準の項目が列挙されている。さらに、チェック項目の総数等も示されている。しかし、具体的な評価法の詳細については何ら記述がなされていなかった。その一方で、順位の根拠とされるスコアは小数点以下2桁まで示されていた。また、ゴーマンレポートの場合、多くの順位はその小数点以下2桁目の数値の違いによって区分されていただけでなく、同じ点数の大学については、何らの根拠を示さずに異なる順位がつけられていたのである。

例えば、1997年に刊行された第10版の場合、トップはプリンストン大学で4.95点であったが、次点はハーバード・アンド・ラドクリフ大学(ラドクリフ・カレッジ)で4.94点、そして、ミシガン大学アンナーバー校が4.93点で第3位となっていた。この他トップ100大学のほとんどは0.01点差で異なる順位が与えられており、0.02点以上で順位に差があったのは10組のみであった(最も多い場合でも0.05点差)。一方、14組については同じ点数であったにもかかわらず順位に差があり、そのうち95位(フロリダ州立大学)から100位(スワスマア大学)までの6校は全て3.91点で同点であった。(ゴーマン 1998: 2-6; Selingo 1997; 小林 2004 をも参照²)。

(2) 症例2：就職力ランキング——小数点以下2桁で順位の明暗

ゴーマンレポートの場合に限らず、大学ランキングに関する以上のような問題点はこれまで何度となく指摘されてきた(Hazelkorn 2015; 石川 2016; Yudkevich 2016)。しかしながら、各種の大学ランキングでは未だに小数点以下の数値で順位がつけられている例が少なくない。この点に関しては、日本で発表されてきた各種の大学ランキングも例外ではな

2 第10版のゴーマンレポートには世界の主要大学のランキングも含まれていた。そのランキングでもやはり0.01点差で順位が分けられている場合が少なくない。ちなみに、日本の大学は「教授の質」と「学問の質」で東京大学がそれぞれ43位と41位にすべり込んでおり、カリキュラムの質に関しては日本の大学のランクインは皆無であった(ゴーマン 1998: 308-313)。以上の点からすれば、ゴーマンレポートの順位というのは、日本から出品された食品(中には納豆や味噌だれというものもある)の8割が入賞するという「モンドセレクション」の評価と類似した点があるのかも知れない。経済法令研究会(2012)参照。

い。例えば、2019年に発表された大学の「就職力」に関するランキングは、その典型例の1つとしてあげられる(『日経 CareerMagazine 価値ある大学2020年版就職力ランキング³』)。そのランキングでは、主に企業の人事担当者に対する「イメージ調査」(回答企業数815社、回答率17.1%)にもとづいて、「学生のイメージ」と「[教育や研究に関する]大学の取り組みへのイメージ」の評価点を集計した上で各大学を順位づけしている。その冒頭に掲げられていたのが図1である。

図1 0.01点差で明暗が分かれる大学ランキング

総合ランキング						
順位	大学名	総合得点	就職力 イメージ	学生の イメージ	教育・研究 のイメージ	社会貢献 のイメージ
1位	九州大学	33.41	8.78	7.79	8.82	8.02
2位	京都大学	32.78	8.13	7.66	9.25	7.74
3位	大阪大学	32.16	8.17	7.69	8.78	7.52
4位	東北大学	32.15	7.75	7.36	9.15	7.89
5位	広島大学	32.06	8.49	8.14	8.40	7.03
6位	宇都宮大学	31.88	8.05	7.93	8.38	7.52
7位	横浜国立大学	31.87	8.02	7.99	8.48	7.38
8位	筑波大学	31.63	7.85	7.52	8.53	7.73
8位	東京海洋大学	31.63	7.98	7.89	7.92	7.84
10位	大阪府立大学	31.51	8.06	8.03	8.38	7.04

出所：日経 CareerMagazine (2019: 7) に矩形を加筆

このランキングで注目に値するのは、ゴーマン・レポートの場合と同じように0.01点という僅差で順位の「明暗が分かれて」いる例が幾つか見られる、という点である。例えば、図1の総合ランキングで言えば、3位の大阪大学が32.16点だったのに対して4位の東北大学は32.15点であった。また、宇都宮大学(31.88点、6位)と横浜国立大学(31.87点、7位)もやはり0.01点差で順位が違っている。しかも、この図から見てとれるように、上位3位までは大きめの活字でしかも大学名の前に金銀銅のオリンピックメダ

3 この冊子の裏表紙には本学のオープンキャンパスの広告があった。

4 点数の算出方法自体にもかなり疑問がある。この調査では、就職力ランキングを上場企業と「一部有力未上場企業」の全4779社(回答率は17.1%、回答企業数は815社)を対象にして「学生のイメージ」を評価してもらった結果を「就職力」の指標としているのである。詳しくは、日経 CareerMagazine (2019: 6) 参照。なお、某大手教育情報関連企業の系列の研究所がおこなっていた「大学満足度調査」に関しては、質問票が非開示となっていた(佐藤2015: 207-209 参照)。その理由について質問してみたところ、その回答は「過去に何度か質問票の無断盗用」があったから、という要領を得ないものであった(回答の受領は2010年4月)。また、その後の追加の質問に対しては回答をいただけなかった。

ルを模したと思われるマークが付けられていた。一方、4位以下は「無印」であり、しかも小さめの活字で大学名があげられている。同じように、全国を7ブロックに分けた「エリア別ランキング」については、0.01点差で順位に差がつけられていたケースが5例あった。

一方で、各種の大学ランキングの中でも最もよく知られている *Times Higher Education* の世界大学ランキングの場合は、小数点以下1桁で順位が決められている。例えば、2018年版の同ランキングで10位に入ったペンシルベニア大学とイエール大学、18位に入ったカリフォルニア大学バークレー校と19位のコーネル大学の場合の得点差は、それぞれ0.1点であった。同じように、このランキングでは、0.2点差で順位が異なる例も多数見られる (*Times Higher Education* 2017: 21)。小数点以下が2桁になるか1桁になるかの違いは一応あるが、大学のランキングに関して、果たしてこのような得点差に意味があるかどうかは意見の分かれるところであろう。

また、大学全体を対象とするランキングそれ自体に果たしてどのような意味があるかについても議論が分かれるところであろう。この点については、少し長くなってしまうが、物理学者である須藤靖による以下の指摘を引用しておきたい。

そもそも政治的・文化的背景を全く異にする世界各国の大学につけられた順位を真に受ける人間がいること自体信じがたい。……かなり度を越した厳密な入試が行われている日本国内ですら、偏差値による大学の序列化は学部あるいは学科別になされている。法学部、文学部から工学部や医学部にわたる偏差値を平均して大学入学難易度ランキングを作成したところで、受験生にとつては何の役にも立たない。もしそんな無意味な統計を発表するレベルの予備校があるとすれば、高い月謝を払ってまでそこに通う意味はないと忠告したい。

同様に、タイムズ・ハイヤー・エデュケーション (THE) とやらが発表する世界大学ランキングの結果を真剣に受けとめる善男善女がいたとすれば、お近くの大学で学び直して、世の嘘と誠を見分ける力を身につけることをお勧めしたい (須藤 2016: 30)。

ここで須藤が述べているように、いかに小数点以下の数値が示されていとしても、そもそも学部等ではなく大学全体をひっくり返して順位をつけるというのは「無意味な統計」以外の何物でもない (同様の点が、大学の全体ないし各学部等を通した平均的な「就職力」に関しても指摘できることは言うまでもない)。また日本も含めて世界中の大学と大学人が大学ランキングだけでなくジャーナル・ランキングなどに狂奔しているという事実から考えれば、たとえもし「大学で学び直し」たとしても、それで数値の「嘘と誠」が見分けられるようになるかは保証の限りではないだろう (トゥーリッシュ 2022: 2章参照)。

いずれにせよ、以上で見てきた大学ランキングに関わる3つの例は、小数点以下の桁を含む測定数値ないし評価数値の精度に関するある種の「信仰」がいかに根強いものであることを示していると言える。

2. ビジネス書・通俗書

(1) 症例 3:『ビジョナリーカンパニー』——整数の合計が小数点以下 2 桁?

上で見てきた大学ランキングの場合とはやや異なる形で測定や分析に関して、「高い精度の測定」という外見を取りつくろっている例もある。その典型が表 1 である。

表 1 『ビジョナリー・カンパニー』に関連する調査結果 (一部)

ビジョナリー・カンパニー	A	B	C	得点	差	得点	A	B	C	比較対象企業
3M	M	H	H	8.00	2.00	6.00	M	M	M	ノートン
アメリカン・エクスプレス	L	M	M	5.00	0.00	5.00	L	M	M	ウェルズ・ファーゴ
ボーイング	M	M	H	7.00	2.00	5.00	M	L	M	マクダネル・ダグラス
シティコープ	M	M	H	7.00	0.00	7.00	M	M	H	チェース・マンハッタン
フォード	M	M	H	7.00	0.00	5.00	M	M	H	GM
GE	H	H	H	8.00	4.00	4.00	L	L	M	ウェスチングハウス
ヒューレット・パッカード	H	H	H	9.00	2.00	7.00	H	M	M	テキサス・インスツルメンツ
IBM	H	H	H	9.00	5.00	4.00	L	L	M	パローズ
ジョンソン&ジョンソン	H	M	M	8.00	2.00	6.00	M	M	M	ブリストル・マイヤーズ
マリOTT	H	H	H	9.00	3.00	6.00	M	M	M	ハワード・ジョンソン

(以下略)

出所: コリンズ&ボラス (山岡訳) (1995: 428)

これは、1994 年に原著が刊行されて世界的なベストセラーになった『ビジョナリーカンパニー (原題 *Built to Last*)』の巻末にあげられていた、同書における議論の根拠とされる調査の方法論に関する付録の一部である。

同書の解説によれば、この表は 18 社のビジョナリー・カンパニー (表の左側) およびそれぞれの比較対象企業と見なされた 18 社 (表の右側) の合計 36 社における「カルト的な企業文化」の程度を各種の資料にもとづいて評価して得点化したものだという。評価の中心になったのは、「社員の愛社精神と基本理念への献身を強烈なものにしようとしているか、社員の行動を基本理念に合致したものにしようとしているか」という点である。

表で A, B, C というのは、それぞれ「教化への努力」、「同質性の追求」、「エリート主義」の程度という 3 つの評価項目である。評価の結果、高 (H)、中 (M)、低 (L) とされた場合にそれぞれ 3 点、2 点、1 点が与えられている。「得点」欄に示されているのはその合計である。また、「差」の欄には、ビジョナリー・カンパニーの得点から比較対象企業の得点を差し引いた値が記入されている (フォードと GM の得点差は恐らく原著の誤植であろう)。

数値データの扱い方をめぐる問題について指摘する上で特に必要ではないので、表 2 では原著の表のうち 11 社分のみをあげている。しかし、この部分を見るだけでもビジョナリー・カンパニーとされた企業の方が全体としては「得点」が高く、したがって、著者たちがそれらの企業の方がカルト的な文化の度合いが高いと見なす上での根拠にして

いることが窺えるだろう。

しかし、そのカルト度に関する「得点」の数値は、ごく単純な整数同士の加算によるもののはずなのである。例えば表の一番上の3Mの場合は、 $M+H+H$ 、つまり $3+2+2=8$ となる。同じようにノートンは $M+M+M$ で $2+2+2=6$ であり、したがって3Mとの差は $8-6=2$ となる。しかし、この表では奇妙なことに、これら3つの値がそれぞれ、8.00, 6.00, 2.00と表記されているのである。他の8組16社についても一貫して小数点以下2桁までの数値が記載されているのだが、『ビジョナリー・カンパニー』には、そのようなかなり変則的と思える表記法に関する説明は一切見当たらない。

もしかしたら、これは、方法論に関する付録の他の部分であげられている表の記載の仕方に合わせたのかも知れない。例えば、CEOの平均在籍年数や特定のCEOの在籍期間中の平均株主資本利益率などは、やはり小数点以下2桁まで示されている。しかし、単なる整数値同士の「足し算と引き算」に過ぎない数値をこのような「精度」で示す必要など特にあるはずもない。また、小数点以下の数値が全て「.00」として表記されていることによって、結果としてこの表は「読者フレンドリー」とはとうてい言い難い、かなり読み取りにくいものになってしまっている⁵。

(2) 症例4：『下流社会』——1人でも8.3パーセント？

小数点以下の桁数が2桁以上ではなく1桁であったとしても、混乱や誤解を招きかねないケースは少なくない。その典型が表2である。

表2 団塊ジュニア男性 階層別意識別「幸せを感じるときは、どんなときか」主な項目 (%)

	上	中	下
n	12	40	48
おいしものを食べたとき	83.3	55.0	56.3
家族でいるとき	58.3	37.5	20.8
仲のよい友だちといるとき	41.7	17.5	29.2
子供といるとき	33.3	20.0	6.3
妻と二人でいるとき	33.3	30.0	4.2
体をめいっぱい動かしているとき	41.7	27.5	10.4
ゆっくり休んでいるとき	58.3	55.0	66.7
ひとりでいるとき	8.3	17.5	27.1

出所：三浦（2005: 189）

5 小数点以下の数値の表示という点以外にも『ビジョナリー・カンパニー』については、社史等の資料の読み取りと解釈を初めとする実証的根拠という点で実に多くの問題があるが、ここではそれらの点に関する解説は省略する。詳しくは、例えばローゼンツワイク（2008）を参照。

これは、前稿でも言及した『下流社会——新たな階層集団の出現』(三浦 2005)に掲載されていたおびただしい数にのぼる表の中の 1 点である。同書は、主として、「1 都 3 県在住」の回答者を対象にして 2004 年と 2005 年に実施された 3 種類の質問表調査のデータを元にして書かれている。回答者は合計で 3500 名前後であったとされ、その調査結果の多くが図や表の形式で報告されている。

同書の第 6 章のタイトルは「『下流』の男性はひきこもり、女性は歌って踊る」というものであり、同章では「階層意識別に団塊ジュニア世代の趣味と消費を見ていく」のだとしている。その章には、1971 年から 75 年生まれの「団塊ジュニア世代」に属する(と言われる)調査当時 20 歳代後半から 30 歳代前半の男女それぞれ 100 名の回答が、上・中・下の階層意識別に集計した結果として、合計で 17 点の表があげられている。なお『下流〜』では、この表と同じように幾つかのセルに網掛けがなされている。明記されてはいないものの、どうやらこれは特筆すべきデータとして著者が判断したものを示しているようである。

第 6 章の総ページ数は 26 ページであり、それぞれの表に関する説明は非常に短い。表 2 についても、以下の約 150 字の解説がその全てである。

……「あなたが幸せを感じる時」を選択肢から選んでもらったところ、団塊ジュニア男性の「上」では、おいしいものを食べたとき、家族でいるとき、仲のよい友だちといるときなどが多いが、「下」は、ひとりでいるとき、ゆっくり休んでいるときが多く、体をめいっぱい動かした時などが少ないなど、やはりオタク的な非活動性が特徴である」(三浦 2005: 190)。

『下流社会』には、前稿で主に取りあげた『下流社会 第 2 章』に関して指摘したのと同様に、調査方法に関する問題が相当数含まれている。例えば尺度構成という点に関して言えば、団塊ジュニア男性の「オタク的な非活動性」を表 2 に挙げられているような質問項目で適切に測定できるとはとうてい言い難い。また、回答者がたかだか 100 名ではサンプルサイズという点でも不適切である(詳しくは、佐藤 (2015: 268-269) 参照)。もともと、本稿における議論との関係においてここで特に注目したいのは、比率(パーセント)を示す数値の使い方である。

表 2 では小数点以下 1 桁までの数値で比率が示されているが、上で述べたように、サンプルサイズ自体が 100 に過ぎない。しかも、その 100 名の回答が「階層意識別」に 3 分割されているために、それぞれのカテゴリーに対応する回答者の実数はきわめて少ない。さいわい表 2 には各カテゴリーの回答者数(n)も明記されているので、それを元にして実数を計算してみると表 3 のようになる⁶。

6 『下流社会』は回答者数が明記されている点に限れば良心的(「まだマシ」と言えるかも知れない。ビジネスレポートや通俗書の中には、実数を明記せずにパーセントの数値だけを挙げている例も稀ではな

表3 パーセントを実数にしてみると……

	上	中	下
n	12	40	48
おいしいものを食べたとき	10	22	27
家族でいるとき	7	15	10
仲のよい友だちといるとき	5	7	14
子供といるとき	4	8	3
妻と二人でいるとき	4	12	2
体をめいっぱい動かしているとき	5	11	5
ゆっくり休んでいるとき	7	22	32
ひとりでいるとき	1	7	13

出所：三浦（2005: 189）を元に改変

この表を見ると、改めてそれぞれのセルの数値がきわめて小さいものでしかなく、⁷ 分かる。例えば、先の引用では、「『上』では、おいしいものを食べたとき、家族でいるとき、仲のよい友だちといるときなどが多い」とされているが、それぞれ、10人、7人、5人がそのように回答しているだけである。さらに、表2では回答者の8.3パーセントが「ひとりでいるとき」を選択したという風にも見えるが、これは実は1名——つまり、文字通り「ひとり」——に過ぎない。

一般にサンプリングに関しては、意図している分析にとって適切と考えられるだけの十分なサイズの標本集団を母集団から抽出するだけでなく、〈その標本集団をさらにどれだけの数のサブグループに分けて比較しながら分析を進めていくか〉という点が最も重要なポイントの1つになる。例えば、クロス集計などによって標本集団を「小分け」にしていった結果として、特定のグループに含まれる事例数があまりにも小さなものになってしまった場合は、グループ同士を相互に比較する意味がほとんど無くなってしま⁸う。これについては、経験則として、サブグループの標本数は主要なもので最低100、さらに下位の分類でも20ないし50以上になるように配慮すべきだとする見解がある（佐藤 2015: 222）。

『下流社会』の著者自身の意図がどのようなものであったかは知る由もない。しかし、

ゝ ない。もっとも『下流社会第2章』の場合には、実数の割り出しが意図的に困難にされていると思われる「帯グラフ」も何点かあげられている（三浦 2007: 7章）。

7 パーセントおよびそれに替わる比率の表示法に関する基本的な考え方については、ザイゼン（2005: 15-25）参照。なお、ザイゼンは、サーベいを繰り返しおこなうような場合には、比較のために小数点以下の数字は残しておいた方がよい、ともしている（ザイゼン 2005: 19）。しかし、『下流社会』は明らかに一種のワンショット・サーベイである。また、尺度構成という点で本質的な問題があるこのようなサーベいを繰り返し実施する意義があるとは思えない。

8 Sudman（1976: 30）。

小数点以下の数値を記載したパーセントの数値は、サンプルサイズの大きさ（小ささ）を分かりにくくしたり測定の精度を誇張して示したりするために使用される例が少なくない。実際、そもそもパーセントは「百分率」なのである。したがって、表2のように、サブグループのサイズが12程度に過ぎないデータの内訳を示すために使用するのには、決して適切なやり方だとは言えない。

以上の点に関しては、社会統計学の古典的な入門書である、ハンス・ザイゼルの『数字で語る (Say it with Figures)』における次のような指摘についてもう一度確認してみる必要があるだろう。

少し考えただけでは、パーセントの数値についてはできるだけ正確に計算して、またその結果をそのまま表示すれば、その分だけ本来の目的を達成できるように思えるかも知れない。しかしながら、小数点以下の桁数が1桁分だけ増える度に、パーセントの数値の単純明快さと読み取り易さが損なわれていく。実際、小数点以下の桁数を増やし過ぎてしまうと、もともとの実数よりもさらに読み取りにくくなってしまふ。小数点以下の桁数はパーセントの数値が本来持っている機能にとって障害になる可能性があるもので、その扱いには慎重を期すべきである。(中略)

つまり、[一般的なルールは] 小数点以下の桁が特別の機能を持たないような場合は省略すべきだ、というものである。小数点以下の桁を省略すると格段に見やすくなるし、一方ではいかにも正確さが保証されているかのような誤った印象を与えるのを避けることができるようになる (ザイゼル 2005: 18-19)。

以上の引用は比率（パーセント）に関わるものであるが、係数値等についても同様の点が指摘できることは言うまでもない。つまり、大切なのは、見かけだけの「正確さ」ないし精度などではなく効果的に「数字で語る」、つまり、数値データによって円滑で効果的（で可能な限り実際に正確）なコミュニケーションを図っていく、ということなのである。

3. 学術文献

先にあげた大学ランキングは一般の読者を対象にしたものである。また、『ビジョナリーカンパニー』や『下流社会』は、何らかの調査データにもとづいているとはいえ、本質的には読み物的なビジネス書ないし通俗書としての性格を持つ文献である。これらの文献では、上で指摘したような小数点以下の数値の表記に関する問題点が看過されることも多い。一方で、専門的な読者を想定して書かれる学術的な文献では、通常は、編集委員、査読者、(学部学生や大学院生の場合には) 指導教員等の指導によってあらかじめそのような問題が解決されている場合が多い。しかし学術文献の中には、そのようなゲートキーパーによるチェックを経ずに、あるいはそれを「すり抜けて」しまうことによって、不必要な小数点以下の数値がそのまま公表される例も見受けられる。

(1) 症例5：学位論文

某ファミリーレストランチェーンを対象にして売上高を従属変数（目的変数）、世帯数、客席数、広告費を独立変数（説明変数）として重回帰分析をおこなった。その結果、重回帰係数は0.90042112、決定係数は0.71613729となった。分散分析では F 値が0.01372976となり有意水準5%で統計的に有意であった。また、各説明変数の売上高に対する影響について見ると、 p 値は以下の通りであり、広告費のみが1パーセント水準で統計的に有意であった——0.6163226（世帯数）、0.42931425（客席数）、0.00855797（広告費）。

学位論文の指導をしていると、草稿段階で以上のような文章が提出されることが珍しくない。いずれの数値も小数点以下の桁数が比較的多い——重回帰係数と決定係数については8桁、 F 値の場合は9桁、 p 値は7桁および8桁。それぞれExcelの出力結果をそのまま転載したものである。

これでは、あまりにも数字が細かすぎて読みづらいし、著者が何を主張したいのかもすぐには分からない。もっとも、「やっている感」の演出が書き手の意図だとしたら、それは一定の成功を収めていると言えるだろう。

一方で、より「読み手フレンドリー」の記述を目指すのであれば、例えば、以下のよう書き替えるべきであろう。

某ファミリーレストランチェーンを対象にして売上高を従属変数、世帯数、客席数、広告費を独立変数として重回帰分析をおこなった。その結果、重回帰係数は0.900、決定係数は0.716となった。分散分析では F 値が0.014となり有意水準5%で統計的に有意であった。また、各説明変数の売上高に対する影響について見ると p 値は以下の通りであり、広告費のみが1パーセント水準で統計的に有意であった——0.616（世帯数）、0.429（客席数）、0.009（広告費）

現在では、データセットさえあれば、このような計算や分析はExcelを使って数回のクリックで立ちどころに出来てしまう。かつては、手計算ないし電卓で数時間をかけておこなわなければならなかった計算も今ではほんの数分でできてしまうのである。また、字数には特に制限が設けられていない学位論文の場合には、Excelで出力される表をそのまま論文に「コピー&ペースト」して、本文における解説はごく短いもので済ませている例もある。その場合、読み手の負担はさらに重くなってしまう。当然ではあるが、そのように出力結果を「コピペ」しているだけでは、とうてい書き手が十分に考えた末に発表した分析結果だとは思えない。

同様の点は、紀要や学術ジャーナルに発表される論文について言える場合も多い。実際、その種の文献では、データだけを替えて先行研究で採用されている分析方法を適用している例が少なくない。また、解析法を十分に理解することなくプッシュボタン式統

9 先行研究を批判的に検討することなく、そのテンプレートを丸写しにするだけでは「再現研究」とは言えないだろう。

計ソフトウェアを利用している例も見られる（しかし、少なくとも業績の量産という効能はあるだろう。）

なお、上の例は寺島ほか（2022: 168-170）の技法書に掲載されていた数値例を元にして
いるのだが、原著で出力結果として示されている F 値および p 値と本稿の著者が Excel
を使って得た結果とでは、小数点以下の桁数が異なっている¹⁰。これは、使用した
Excel のバージョンの違いによるものとも考えられる。寺島らは、「執筆時点で最新の
Office 365 を用いた」（寺嶋ほか 2022: (3)）としているが、本稿では Office 2021 for Mac を
使用している。このようなソフトウェアやそのバージョンによる有効数字の違いも、も
し「精度」にこだわるのであれば本来は注意を払うべきポイントの 1 つであろう。ま
た、それは Excel に「使われる」のではなく Excel を「使いこなして」いこうとする際
には留意すべき点でもある。

(2) 症例 6：学術論文における p 値依存症（アステリスク症候群）

上記の例とは少し違った意味で、学術論文や研究書などで「コンマ以下」の数値が必
要以上に強調されてしまう典型的な例に p 値の有意水準というものがある。

上の例では Excel で出力された p 値自体が示されているが、統計的調査では p 値に
ついては、0.05（5%）以下から 0.001（0.1%）を統計的な「有意水準」の目安として設定
する人が多い。また、学術論文に掲載される統計データの分析結果を示す表の数値に
は、1 個から 3 個までのアステリスクが添えられている例が稀ではない。その種の表の
下方には「*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ 」などという表記がよく見られる。そ
して研究者の中には、未だにその星（アステリスク）の数が多ければ多いほど「研究仮説
が検証された」ことの確実な証拠であると思い込んでしまっている人々もいる。

この p 値に関する考え方については、既に 60 年以上も前からさまざまな問題が指摘
されてきた。日本でも、動物心理学者の橘敏明は 1986 年に『医学・教育学・心理学に
みられる統計的検定の誤用と弊害』というタイトルの書籍を著し、統計的有意性の検定
をめぐるさまざまな問題について警告を発していた。例えば、橘は次のように論じてい
る。

したがって、有意性検定の結果で示される「有意な差」の出現というものが客観的なものではありません、
研究者側の操作で決定できる部分があるということである。だから、有意性検定の結果で示される「有
意な差」は、研究者の意図や希望とは無関係の、客観的で科学的な指標であるという一般的にひろく信
じられている考えは覆ることになる。（中略）すでに何人もの研究者によって指摘されているように（例

10 原著では F 値については小数点以下 7 桁、 p 値は 9 桁までとなっている。さらに細かな点について言
えば、有効桁数を正確に示すという点からして、世帯数に関わる p 値は、0.6163226 ではなく客席数お
よび広告数と同じ 8 桁に揃えて 0.61632260 とするべきであろう。なお、原著ではここに示したような
文章による解説がなされているわけではなく、適宜小数第 4 位以下が四捨五入されている。

えば Yetes, 1954) [原文ママ。実際には Yates (1964) である。なお、『医学〜』の文献表には正確な記載がある], 有意性検定の結果に不当に注意がはられすぎて, その結果, 研究している処理効果の実質的な大きさへの評価への注意がおろそかになってしまっているのである。この“有意”という言葉自体が実質的な処理効果の存在をほのめかしているの、よけいに誤解をうけやすい(橘 1986: 72-73) (下線は引用者)。

上に引用した文章の下線を引いた箇所橘は, p 値に関する基本的な誤解, つまり, 〈 p 値が客観的で科学的な指標である〉という考えは「覆る」としている。しかし, この種の基本的な誤解ないし「 p 値依存症」(アステリスクの数に一喜一憂するという点では, 「アステリスク症候群」と呼ぶこともできるだろう)とでも呼ぶべき傾向は, その後 30 年以上の歳月を経て今なお日本のみならず世界のいたるところで健在なようである。さらに, 橘が引用している Yates の論文が刊行された 1964 年の論文の時点からすれば, 60 年近くにわたって同種の誤解が蔓延していたことになる。¹¹それは例えば, 2016 年に米国統計協会から出された「 p 値に関する米国統計協会の声明: コンテキスト, プロセス, そして目的」という声明文で, その著者たちが上の橘の主張とほぼ同様の指摘を繰り返さなければならなかったという事実からも明らかであろう (Wasserstein and Lazar 2016; 佐藤俊哉 2017; ラインハート 2017 をも参照)。

さらに近年では, 橘が指摘していた p 値が持つ「研究者の意図や希望とは無関係の, 客観的で科学的な指標」としての印象はむしろ強まってきているとさえ言えるのかも知れない。何しろ, Excel をはじめとする表計算ソフトや統計解析ソフトを使えば, p 値は小数点以下 8 桁ないし 9 桁の値までほぼ自動的に出力されてしまうのである。

(3) 症例 7: p 値ハッキング (p 値の捻り出し)

もっとも, 以上のような p 値依存症に陥って「統計的に有意」な分析結果に一喜一憂している研究者は, ある意味ではまだ良心的なのかもしれない。というのも, 研究者の中には, 統計的な有意性の限界について十分に承知しながらも, 「有意な」 p 値を強引に捻り出す(偽装・捏造する)人々が少なからず存在しているからである。

その種の, 確信犯とでも呼ぶべき人々の常習的な手口の中でも極端なもの 1 つに, p 値の特定の桁数を四捨五入して有意水準を偽装してしまうというものがある。例えば, Excel ないし何らかの統計解析ソフトウェアで出力されてきた p 値が 0.05432109

11 なお Yates が 1964 年に発表した論文は, その 2 年前の 1962 年に物故したロナルド・フィッシャーを追悼して編集された *Biometrics* 誌の特集号に収録されていたものである。

12 「確信犯」的な行為には少なくとも 2 つのタイプがあるだろう。つまり, 1 つのタイプは, 本節で解説しているような明確な偽装行為に手を染めるケースである。他方には, 統計的な有意性の検定の限界を十分に理解しながらも, 有意性検定を論文の主張の中心に置くようなスタンスを取るというような場合もある。

というものであったとしよう。「確信犯」の場合には、この p 値の小数点以下第3位を四捨五入して0.05とした上で、さらに不等号まで誤魔化して $p < 0.05$ 、つまり「5%水準で有意である」という風に報告するのである(トゥーリッシュ 2022: 137)。

それ以外にも例えば、調査や研究を進めていく中で統計的に有意な結果が出た段階でそれ以上のデータ収集を取り止める、という「手口」はさまざまな学問分野で広く採用されてきた。また、収集したデータの分析の際に、とにかく何らかの点において有意性が確認できるまでさまざまな方法で「データ・マイニング」と検定を繰り返す「データ浚渫 (data dredging)」と呼ばれるテクニックもある。

表4は、一般に p 値ハッキングとも呼ばれる、その種の「データいじり」に関する幾つかの実態調査の例を挙げたものである。この表を見るだけでも、その種の確信犯的な偽装行為がいかに広い範囲に及ぶものであるかが確認できるだろう。

表4 p 値偽装に関する実態調査の例

QPR (疑わしい 行為)	出所	方法 (特記しない限り 質問表調査), サンプル, 国	結果
p 値ハッキング	Boulbes et al. 2018	大学院生とポストドク研究員 467名: 米国	24.2% が作業仮説を支持しない研究結果を省略したと回答
	Fraser et al. 2018	生態学者 494名, 進化生物学者 313名: 各国	27.3% と進化生物学者の 17.5% が p 値の「四捨五入」を認めた
	Tijdkink et al. 2016	生物医学科学者 535名: ヨーロッパ	25% が、分析後に少なくとも1回、データを選択的に削除または報告と回答。60% がデータポイントを意図的に削除した同僚を知っていると回答
	John et al. 2012	心理学者 2155名: 米国	63.4% が従属変数の測定値の全ては報告しなかった。55.9% は、結果が有意であるかどうかを確認した後で、さらに多くのデータを収集した。27.7% が一部の結果だけを報告した。22% が p 値を四捨五入した。45.8% が、「うまくいった」研究だけを選択的に回答。特定のデータの除外による影響を調べた後で 38.2% が実際に除外
	Martinson et al. 2005	科学者 1768名, 博士課程の学生 126名: 米国	15.3% が、観測値またはデータポイントの削除を回答

出所: トゥーリッシュ (2022: 154) 所収の表より抜粋

表4では、主として自然科学および心理学系の論文の例があげられている。しかし、経営学系の研究論文もこのような傾向から決して自由であるわけではない(トゥーリッシュ 2022: 150 参照)。2018年に発表されたある文献レビューによれば、「経営学系の論文のうち25パーセントから50パーセントには何らかの矛盾ないし明らかな誤りがある……全体として、学術研究の真実性および(当然であるが)その結論と含意について疑うだけの十分な理由があることを示すエビデンスが増加しつつある」とされている(トゥーリッシュ 2022: 140-141)。

こうしてみると、数値データやその分析結果の精度ないし科学性に関する執着は過信症などというレベルを通り越して、もはやほとんど治療不能な「宿痼」とさえ言える強

迫観念になっていると言えそうである。¹³

II 測定精度過信症と「やっている感」の受容（診察②：What）

1. 「過信」しているのは誰なのか？

以上で検討してきた幾つかの事例は、小数点以下の桁数については、1桁（『下流社会』）、2桁（大学ランキング、『ビジョナリー・カンパニー』）、7桁から9桁（重回帰分析）と、見かけ上は大きな違いがある。また、数値の性格も単純な比率（パーセント）（『下流社会』）、評価点（大学ランキング、『ビジョナリー・カンパニー』）、統計的検定の結果（重回帰分析）と多様であり、必ずしも「測定値」という言葉でひと括りにすることはできない。

もっとも、一方でそれら複数の事例（症例）には1つ重要な共通点がある。それは、小数点以下の数値によって、「社会現象を正確かつきわめて厳密に測定した作業の成果」という印象が醸し出されているという点である。そして多くの場合、それは、単なる見せかけに過ぎない。また、重回帰分析の係数値のように、多すぎる桁数がむしろ円滑なコミュニケーションにとって障害である夾雑物になっている場合も多い。そのような面があるからこそ、本稿ではこれらの事例を測定精度に関する過信症の「症例」と呼んできたのである。

もっとも、ここで精度について「過信」しているのは必ずしも書き手であるとは限らない。実際、その書き手たちは、読み手の側が、コンマ以下の細かな情報を含む数値を科学的で客観的な測定の結果として受け入れる可能性があると想定しているからこそ、そうしていたのだと考えられる。あるいは書き手の方は、それが単なる見かけに過ぎないことは十分に承知しており、むしろ数値の精度に関する読み手の側の思い込みを利用して、いわば確信犯的に不必要な精度を含む数値情報を提示していただけたのかも知れない。その場合は、書き手は「コンマ以下」の精度などについては最初から信じてなどおらず、むしろもっぱら読み手の方が「過信」している可能性があると言えるだろう。

2. 受容の実例——「データに実証性」、「しっかりしたデータ」

例えば、先に述べたように、ゴーマンレポートは、1967年に発刊されてから1997年までの30年間に10版を重ね、また、学部教育課程だけでなく大学院課程と専門職スクールのランキングに特化したレポートも刊行されていた。またゴーマンレポートとは違って順位を決定する上での根拠となる指標や順位の元になる得点の算出方法は一応公

13 また、 p 値ハッキングに手を染めていない論文の場合でも、肝心の仮説命題自体が単に「ある変数の値が増加すると、別の変数の値が増加または減少する」と述べているだけに過ぎない例が多いともされている（Edwards and Berry 2010: 670; Tourish 2019: 141-142）。

開しているものの、その方法について何度となく疑義が提出されてきた *Times Higher Education* によるランキングを典型とする各種の世界大学ランキングは、2000 年代初頭から猛威を振るい、各国の高等教育政策を左右するほどの影響力を持ってきた。そして、その多くは、*Times Higher Education* の場合と同様に小数点以下の数値を含む評点にもとづいて順位を決定しているのである。(Hazelkoron 2015; Shin et al., 2011; Yudkevich et al., 2016; 石川 2016)。

同じように、『ビジョナリー・カンパニー』は 1994 年に刊行されて以来、世界的ベストセラーになった。また同書の著者の 1 人であるジェームズ・コリンズが、(少なくとも日本では) その続編という触れ込みで次々に刊行されてきた経営指南書もそれぞれベストセラーとなっている (コリンズ 2001, 2012, 2017, 2020, 2021)。

以上の例と比べれば影響は日本に限定されるものであるが、『下流社会』には、何人かの大学教員からは書評などの形で賛辞が寄せられていた。その中には、専攻が社会科学系である人々、つまり最低限のリサーチ・リテラシーを身につけており、数値データの扱い方には習熟している (べきだ) と思われる人々も含まれている¹⁴。以下にはその数例をあげておく。

市場調査を土台とするしっかりしたデータにもとづいて、低収入層の生活スタイルを分析している。(野口 2005)

多様な統計的資料、量的調査結果などを駆使し、現代社会の特徴を分析し、いま私たちがどこにいるのか、どのような場所で、どのような生活の層で暮らしているのかを明快に説明してくれる作品 (好井 2006: 9-10)

データに実証性がある、具体例が豊富 (上野) (上野・三浦 2007: 27 における上野の発言)

マーケティングを専門とする著者が、独自の調査データによって、近年の「中の下」の姿を浮きぼりにしている (岩井 2005)

上記の文章に見られる「量的調査結果などを駆使」(好井)、「データに実証性」(上野)、「市場調査を土台とするしっかりとしたデータ」(野口) という評言は、それらの著者のリサーチ・リテラシーについて基本的な疑念を生じさせるものである。もっとも一方でそれらの評言は、数値形式で表現された情報が、(少なくとも日本国内では) 本来ある一定以上の批判能力を持つことが期待されている大学教員も含めて受容されている、という可能性を示唆するものである。言葉を換えて言えば、測定精度過信症は、どうやら

14 それぞれの賛辞が発表された当時の評者の肩書きは、以下の通りである。野口－神戸大学名誉教授，好井－筑波大学教授，上野－東京大学教授，岩井－京都大学教授。

社会全体に広く蔓延しているようなのである。

Ⅲ 不可解なグラフと測定精度過信症の背景 (診断: Why¹⁵)

1. 根底にある数値信仰

(1) 数値データ=現実を曇りなく映す鏡？

本稿で事例(症例)を紹介してきた小数点以下の情報を含む数値の精度に対する思い込みないし執着の根底には、物理現象を測定するために使用される計器から出力されてくる「コンマ以下」の桁を含む数値からの連想があると思われる。また、それらの数値の客観性や科学性に対して我々が寄せている信頼も重要な役割を果たしていると見てよいだろう。実際我々は、熱があるかどうか自信が持てない時には体温計を使う。また、空調機(エアコン)の効き具合について意見が分かれた時には、室温計に表示された温度の数値によって議論に決着がつけられる場合が少なくない。同じように、正確な長さについては物差しをあてて確認するだろうし、正しい重さを知るためには秤(はかり)を使う。

人間の五感による印象は、ともすればきわめて不安定で不確実なものでしかない。それに対して、精密器械は時には人間の感覚器官ではとうてい感知することができないほどの、ミリ単位ないし「ナノ単位」の微細で詳細なデータまで正確かつ精密にかつ安定して記録することができる場合がある。

実際、物理的な現象や状態については、各自の主観的な印象であれこれ議論を重ねるよりは、何らかの測定器具を使って客観的な証拠を示した方がはるかに確実である。つまり、物理的な測定値は議論が紛糾した時に決着をつける上で有力な根拠になり得るのである。実際、それらの数値は、人為的な操作や人間の主観が介在することなく生み出された客観的で正確な情報のように見える。だからこそ、数値データそれ自体あるいはそれを加工してグラフや表として示した場合には、「データそれ自体に語らせる」ことができるように思えてくるのである。

それと同じようなことが、社会的な現象について判断したりその判断に基づいて意思決定したりする際にも言える。事実、経済・経営指標などのような数値情報は、経済政策や経営方針について議論が紛糾した時に、よく「客観的」な科学的根拠(エビデンス)として引き合いに出されることがある。そして、そのような時によく使われる言葉の1つが「数字は嘘をつかない」というものである。

実際、前稿で取りあげたグラフの誤用例や本稿で見てきた小数点以下を含む数値の使用とその受容の背景にあるのは、数値として示される統計的なデータを、社会的な現実の姿を1点の曇りもなく映し出す鏡のようなものと思い込んでしまう傾向であると考え

15 本節および次節の一部は、佐藤(2021: 293-300)を元にしてている。

られる (Alonso and Starr 1983: 1; Starr 1983; 52 参照)。つまり、我々は、ともすれば次のような幻想を抱いてしまいがちなのである——〈小数点以下を含む数値データやそのデータを元にして作成された図ないしグラフというものは、複雑な現実世界を精確かつ効率的に要約して示してくれる魔法のような力を持っている〉(佐藤 2015: 10; 2021: 303)。

(2) 物理的測定 対 社会的測定

しかしながら、実際には、物理的現象に関して用いられてきた測定機器と社会現象の測定のために考案されてきた各種のモノサシ (指標・指数・尺度) とのあいだには、幾つかの点で本質的な違いがある。表 5 は、その両者の違いを「物理的測定 対 社会的測定」¹⁶ という形で整理してみたものである。

表 5 物理的測定 対 社会的測定

	物理的測定の理想	社会的測定の現実
国際的な標準器	有	無
ヒトの関与 (判断・選択・解釈等)	無	有
現実形成力	無	有
時代・社会の変化による影響	無	有

出所：佐藤 (2021: 294)

まず、物理的な測定の場合には、例えばメートル原器 (長さ)、キログラム原器 (重さ) などのように究極の基準になる国際的な標準器が 100 年以上にもわたって使われてきた。それに対して、社会的な測定については、必ずしも国際的に認められた基準が決定的な影響力を持ってきたわけではない。むしろ、無数の理論上の流派が設定した基準や指標が「百花繚乱」のように林立している場合が少なくない。

この基準の林立という例からも分かるように、社会現象に関する測定値の場合は、人間による判断や解釈が関与する余地がかなりある。また、測定に関わる人々の理論的立場や利害関心によって、どのような項目をどれだけの重み付けで考慮に入れるかという判断も大きく左右される。

これは 1 つには、社会現象に関する測定値は、それ自体が 1 個の事実になって社会的現実を変えていく潜在力を持っているからである。例えば、本稿で取りあげた大学ランキングは、まさに現実形成力を持つ数値の典型例だとも言える。同じように、企業を対象にした各種のランキングや格付けは、それ自体が重要な原因になってその企業の株価やブランド価値を左右していくことがある。

以上のように、社会的測定には、永久不変の自然法則の存在を想定してそれを解明す

16 この表の内容については既に他のところで比較的詳しく論じたことがあるので (佐藤 2021)、ここでは、本稿のテーマと関連の深い事柄に焦点を絞ってごく簡単に述べておくことにした。

る努力の一環としておこなわれる物理的測定とは非常に対照的な面がある。これは、社会や経済、経営というものが、測定法の進化や変化をはるかに超えるほどの猛烈なスピードで刻々と変化する、いわば「生き物」のような存在だからでもある。

(3) 0.01 秒 対 0.01 点——社会的測定の精度

以上で見てきたように、外見上は同じ数字のように見えても、物理的な測定値と社会的な測定値とのあいだには本質的な違いがある。

特に重要なのが、測定の精度に関する違いである。物理的な状態や現象に関する測定値の場合は、先にあげた標準器が究極のよりどころとなる国際的な基準や厳密な手続きを通して非常に高い精度が保証されている。例えば、我々は、100 メートル走の日本記録について、2022 年時点の第 1 位は同年 6 月に山県亮太が出した 9.95 秒で、2 位が 9.97 秒のサニブラウン・アブデル・ハキーム (2019 年の記録)、9.98 秒で 3 位に並んでいるのが桐生祥秀 (2017 年) と小池祐貴 (2019 年) である、と言われても何ら違和感は覚えなだろう。

一方で、社会的な測定の場合には、それと同じレベルの精度が保証されることは滅多にない。ところが現実には、本来それだけの精度が保証されないはずの数値が一種の社会的な事実として一人歩きしてしまう例が珍しくないのである。その典型が、先にあげた大学ランキングである。中には、ゴーマンレポートのように、小数点以下 2 桁の数値で順位に差がつけられている例もあり、その精度や信頼性はこれまで何度となく厳しい批判にさらされてきた。一方で、本稿で例としてあげた重回帰分析の結果として算出 (出力) される各種の係数値がそうであるように、「小数点以下の桁数が多ければ多いほどより高い精度が保証される」という思い込みも生じてきたのである。

このように、社会的測定には非常に多くの限界や制約があるにもかかわらず、我々はその測定値について物理的測定と同じような精度や正確さを期待してしまいがちである。その意味では、測定精度過信症をはじめとする社会的測定に対する過度の信頼については、これを「数値信仰」と呼ぶことができるだろう。

2. 見果てぬ夢としての厳密社会科学 (exact social sciences)

(1) 物理学への羨望

このように、社会的な測定に対して物理的な測定と同じような精度を期待するのは「無いものねだり」でしかない。その点からしても、経済学者や経営学者を含む一部の社会学者たちが抱いてきた「物理学への羨望 (physics envy)」が、現実離れた幻想のような願望に過ぎないことが改めて明らかになってくる。

その羨望とは、数学に基礎を置いて目覚ましい成果をあげてきた物理学とその理論

的・実証的方法へのほとんど崇拜と呼んでもいいような憧れの感情に他ならない。その憧れが募るあまりに、物理学と同じような方法で社会現象を精確な数値データと洗練された数式の形で描き出すことを目指そうとするのである。そして、上で指摘した社会一般に広く根付いている数値信仰の背景にも、このような、物理学を典型とする「ハードサイエンス」ないしそこで使用されている（と想定されてきた）「ハードデータ」を絶対視してしまうという思い込みがあると考えられる。つまり、ここでは「エビデンス＝科学的根拠（エビデンス）＝数値データ」という等式のような関係が想定されているのである。

もっとも、もしかしたらそのような物理学への羨望を抱いているのは定量的な方法を駆使する社会科学者のごく一部だけなのかも知れない。実際、最先端の研究に取り組んでいる研究者ほど、数値データによる社会現象の分析が持つ限界については明確に認識しているとも思われる。また、そのような研究者の多くは、現実におこなわれてきた社会的測定には、多かれ少なかれ「ゴム製のモノサシ」のような一面が認められることを熟知しているだろう。当然であるが、ゴムを素材にしたモノサシは、力の入れ加減で伸び縮みするし、目盛りと目盛りのあいだの間隔も不均等になりがちである。

良心的かつ賢明な研究者は、そのような点を念頭に置いて、その限界内で最善を尽くしているものと思われる。一方で、修業中の初心者には、ゴム製のモノサシでしかない社会的な測定結果についてもあたかも測定原器を使った厳密な「校正」を受けた測定器で測った値と同じような精度が保証されるものだと思込め込んでしまうのかも知れない。だからこそ、彼／女らは、「コンマ以下」の桁数にこだわるのであろう。

(2) 厳密社会科学の見果てぬ夢

物理学への羨望と密接な関連を持つ言葉に「厳密科学 (exact science)」というものがある。ここで厳密科学（「精密科学」とも）というのは、「厳密な量的規定による論証体系に組織される科学。数学、物理学、科学など」（小学館『精選版日本国語大辞典』）を指すとされる。¹⁷社会科学においても、それらの厳密科学をモデルにして、社会現象についても、精確な数値データを駆使し、また厳密に定義された概念や洗練された数式の形で表現された法則の体系としての理論による説明を目指そうとする動向が存在していた。この点において急進的であったのは心理学系の「行動科学」や経済学であると言えるだろう。

物理学への羨望と密接に関連する「厳密社会科学」への志向の1つの典型は、社会学者のジョージ・ランドバークが1939年に著した *Foundations of Sociology* (Lundberg 1939) に（ほとんど戯画化された形で）¹⁸ 見ることができる。ランドバークは、社会学における定量

17 『哲学事典』（平凡社 1971年版）および『哲学・思想事典』（岩波書店）をも参照。なお、『精選版日本国語大辞典』および『哲学事典』では「精密科学」の見出しになっているが、『哲学・思想事典』では「厳密学／精密学」となっており、厳密学と精密学を区別すべきだとしている。

18 ランドバークは、コロンビア大学やスタンフォード大学で教鞭をとり、1943年には第33代の「米国ノ

的アプローチの有力な論客であり、また、概念の操作的定義という考え方を社会科学に導入した立役者の1人としても知られている。同書の序文でランドバークは、物理学と比較した場合の社会科学の未発達な状況について指摘した上で、次のように主張している——「もし我々がこの方法を物理学の場合と同じように社会科学についても忠実に従っていけば、それに応じた統制力という果実が得られるであろう」(Lundberg 1939: 20 ivi)。また、同書を手厳しく批判したレスリー・ホワイトは、あるところで同書の基本的なメッセージについて次のように要約している——「物理学に注目しようではないか。物理学は厳密科学であり、驚くほどに効果的で目覚ましい成功を収めている。その成功は主として数学の使用によるものである。我々社会学者も同じような道を進んでいくのではないか」(White 1943: 374²¹)。

物理学の論証の方法と法則体系、そしてまた実証における方法論をモデルにして社会学を厳密科学として鍛え上げようとする試みは、*Foundations of Sociology* では物理学理論(量子力学、場の理論、電子と原子の理論など)やニュートン、ボーア、ハイゼンベルク、アインシュタインなどの物理学者の名前が頻繁に言及されているという事実からも窺える。例えば、同書ではアインシュタインの名前が16回引用されているのに対して、デュルケームは2回のみである(White 1943: 373)。(ウェーバー、マルクス、ジンメルへの言及はそれぞれ3回に過ぎない。)しかしながら、ピティリム・ソローキンによれば、ランドバークが同書において開陳した彼の物理学や数学に関する理解は基本的な誤解を含むものであったとされる(Sorokin 1940)。またランドバーク自身が展開した社会学の基礎に関する議論は、協働、慣習、社会化など旧来の概念とそれほど異なるものではなく、彼がモデルにしようとした物理学の概念体系や法則体系からはほど遠いものでしかなかった(White

ゝ 社会学会会長をつとめている。

19 これについては、Lundberg (1939: 54-56) をも参照。なお、*Foundations of Sociology* では、操作的定義概念のアイディアを含むパーシー・ブリッジマンの *The Logic of Modern Physics* (Bridgman 1927) への言及は4回なされている(Lunderberg 1939: 38, 43, 81, 83)。また、ブリッジマンの名前は合計で9回引用されている。

20 ランドバークは、事例研究に対する否定的な見解でも知られており、あるところでは事例研究と統計的研究は必ずしも対立するものではないとしつつも、次のように述べている——「事例研究法それ自体は科学的方法などではなく、単にその第一歩に過ぎない」、「量的な要約が欠如している中での事例記録にもとづく一般化は科学的研究にとってはほとんど役に立たない」(Lundberg 1926: 61, 62) (ランドバークはこの論文で事例研究法を「統計という風に付けられた頼りないしっぽ(a helpless tail to the statistical kite)」という表現を使ったとされることがある(Denzin 1978: 233; 佐藤 2006: 77)。しかし、実際には同論文にそのような記述は見当たらない。これは、どうやらゴードン・オルポートの論評(Allport 1942: 55)をランドバーク自身の言葉として引用してしまったものらしい。ちなみに、この全体で5ページ、語数にして3300語に満たない論文で science という言葉は10回、scientific は29回使用されている。この点からも、社会学を「科学」にしていこうという当時のランドバークの並々ならぬ意気込みが窺える。なお、事例研究に関してよくある誤解については、Flyvbjerg (2011) 参照。

21 ピティリム・ソローキンも次のように要約している——「ランドバークの主張は、社会学は物理学や厳密科学と同じような方向で再構築すべきであり、またそれらの自然科学の方法を見習うべきなのだ」(Sorokin 1940: 795)。なお、ランドバーク自身が自著の中で exact science(s) という言葉を使用しているわけではない。

1943: 374)。

その後も、例えば行動科学的なアプローチをより本格的な形で導入して社会科学を少しでも厳密科学に類似したものに近づけようという試みは幾つかあった (Ellul 1976)。しかし、それらの試みの一部は統計的分析の洗練という方向で生かされていく一方で、社会現象に関する定量的なアプローチの適用範囲やその限界が次第に明らかになっていくことによって一時の熱狂は明らかに過去のものになっていった。そして、ある時期からはむしろ人文モデルが力を盛り返してきたとさえ言われている。

その点に関しては、社会科学における「解釈学的転回」に関する論考として知られるクリフォード・ギアツによる 1980 年の次のような指摘が最もよく知られているものの 1 つである。

多くの社会科学者が説明における法則と例証という理念から事例と解釈という理念に注意を向けるようになり、また、惑星と振り子を結びつけて考えるような発想から菊と刀の関係に関する発想を重視するようになっている。さらに、社会科学的な現象を理解する際に物理現象の理解に関して長いあいだ工芸や技術のアナロジーが果たしてきた役割を今や人文学の領域に由来するアナロジーが果たしてきている、というのも事実である (Geertz 1983: 19 [初出は 1980 年])

もちろん、社会現象の分析において物理学由来のアナロジーが完全に「お役御免」になったわけでもないし、人文学系の発想によって社会的行為の全てが理解できるわけでもない。²²しかし、少なくとも社会科学を物理学や化学のような厳密科学に変換していこうとする試みの多くは「見果てぬ夢」に終わったと言えるだろう。

(3) データに語らせるという幻想を越えて——「数字が語る」から「数字で語る」へ

もっともその一方で、社会科学の「厳密社会科学化」を目指すかつての熱狂ほどではないにしても、一般社会においてもまた研究者の世界においてもある種の数値信仰は確固たるものとして存在していると見ることができる。それを象徴的に示している例が、先にもふれた「データそれ自身に語らせる」という言い回しである。

「データそれ自身に語らせる」というのは、一般的には「良質のデータさえあれば、社会的な事実(ないし、「真実」)はおのずから明らかにできるものだ」とするような考え

22 実際、ギアツは人文学系のアナロジーこそが最適だと主張しているわけではない。彼は、この引用のすぐ後で次のようにも述べているのである——「社会科学に関して言えば、よく嘆かれるように固有の領域を持たないことが社会科学を特色づけるのではなくなったということである。社会科学は未発達な自然科学であって、時間さえあれば、また進歩した分野からの助けさえあれば、より確固たるものになるはずであると考え、あるいは社会科学は人文科学の使命の無知で自惚れな篡奪者であり、確固たるものなどそもそもありえないところにそんな空約束をするだけであるとみなすこと、あるいはまた社会科学は他の学門領域とは明らかに異なった営みなのであって、『2つの科学革命』で指摘された C. P.] スノーの『自然科学 対 人文学という』規範的な二分割の中間に位置する第三の文化であるとみなすこと等、そうしたことは以前にも増してあり得ない話になってきている」(Geertz 1983: 21)。

方に他ならない。この言い回しやそれに類似した発想それ自体は、例えば現場調査を重視し物事の実態をつぶさに記述することを目指す、いわゆる「現場主義」的な研究者の発言や著作に見られる場合も少なくない。もっとも、何らかのデータそれ自体が何よりの最終的な証拠ないし根拠（エビデンス）となり、それによって議論の黒白がつけられるという考え方は、数値データを主として用いる研究者の場合にはより顕著であるように思われる。

また一般的にも、「数字は嘘をつかない」という言い回しがあるように、文字や音声などのいわゆる定性データにくらべて主観的な解釈の余地が少ないとされる「ハードデータ」、すなわち数値化されたデータに対する信頼は絶大なものがあるだろう。それは例えば、「エビデンスベースト」などという際に、ほとんど自動的にエビデンス＝科学的根拠＝数値データという等式のような理解が成立してしまうことにもつながっている。²³

そして、以上のような場合には、「数字が（おのずから）語る」という想定がなされていると考えることができる。また、だからこそ、測定精度に対する過剰とも思えるほどの信頼が生じてきたのだろう。また、「おのずから語る」ものと見なされてしまった数値データを図解の形で表現したグラフが、社会的な事実を鏡のように映しだしていると考えられてきたのだと言える。つまり、我々が小数点以下の数値やグラフに寄せる信頼は、「数字が（おのずから）語る」あるいは「グラフそれ自身が語る」ないし「グラフそれ自身に語らせることができる」と想定しがちなのである。

しかし当然ではあるが、データそのものは決して語ることはない。人間が特定の枠組みにもとづいて適切な解釈をし、ふさわしいレトリックを用いることによって初めて「データそのものが語る」ように見せかけることができるだけなのである。あるいはまた「データとともに語る」ことができるだけと言ってもよいだろう。「データ自身が語っている」ように見えるのは、知らず識らずのあいだに身につけてしまった特定の物の見方に囚われているために、むしろ「データに欺かれた」時であるとさえ言えるだろう。²⁴

言葉を換えて言えば、データ分析の際に重要なのは、数字それ自体が語ることを受け身の姿勢で待ち受けるのではなく、むしろ我々自身がいかに明快にかつ正確に「数字で語る」ないし「グラフで語る」ことを目指していくことなのである。実際、言葉を駆使して自分の思いをできるだけ正確かつ効果的に伝えようとする際には各種の工夫が必要であるように、数字で語る際にも、さまざまな工夫と習熟が不可欠になる。そのような

23 それがまた数値データにもとづくランダム化比較試験（RCT）を EBPM の究極の理想型としてとらえる、ある種安易な発想に結びついていくことにもなった（Cairney 2016 参照）。

24 以下の文献を参照——Arnold (1970: 149), Blumer (1979: 122), リクルール (2004: 270), 米盛 (2007: 159-160), Singleton and Stait (2010: 88, 418)。

点で不可欠となる基本的な知識の中に含まれているのが、「データの縮約」という考え方である。

3. 拮抗する2つのベクトル——縮約 対 膨張

数値化の作業を通して社会現象を明確かつ明快に表現しようとする作業には必然的にデータの「縮約」などと呼ばれる手続きが含まれている。もっともその一方で、定量的な分析の中には、縮約とは正反対に、むしろデータの構造や意味を分かりにくくしてしまう要素が入り込んでしまうこともある。前稿で検討した図解表現における各種の夾雑物や本稿で見てきた不必要なほど多い小数点以下の桁数などは、本稿で「膨張」と呼ぶ、そのような性格を持つ、本来は不必要であるはずの要素の典型例である。

(1) データの縮約とは？——抽象化と省略によるモデルの形成

数字やグラフによって社会現象について分かりやすく表現して（「語って」）いこうとする際に不可欠になってくるのが、データの縮約（data reduction）という手続きである。データの縮約というのは、収集された大量の数値データの集合を何らかの操作を加えて整理することによって、秩序や規則性、ないし複数のデータの関係性などを見出していく作業である。つまり、それだけでは一見混沌とした数値の塊のようにしか見えないものの中から構造や意味を見つけ出していく作業である。

縮約の代表的な手続きの1つに、平均値や分散などのいわゆる要約統計量（summary statistics）を算出するというものがある。これは、膨大なデータ全体の特徴を1個ないし数個の単純な数値に集約ないし凝縮して示す、というアプローチである。また、集計表の形で数値データ全体の関係性ないし構造を示したり、グラフとして図解して視覚的に表現する手続きもデータの縮約の一種であるとされることが多い（Bailey 1994: 338; Smithson 2000: 21-22; 小島 2006: 17-18; Bryman and Bell 2015: 13）。

このデータの縮約という作業は、通常は数値データの分析の手続きとしてとらえられる。もっとも、文字テキストを典型とする定性的データをコーディングして共通のテーマを割り出していくような作業も縮約プロセスの一種として考えることができる（Miles and Huberman 1994: 10-11; Bryman and Bell 2015: 13; 佐藤 2008; クカーツ 2018）。実際、例えば文字起こしたインタビューの記録などのように、一見混沌とした無秩序な文字の塊のようなものにしか見えないテキストデータは、コーディングなどの作業を経ることによって、その構造と意味がより明確な形で浮かびあがってくることが多い。

以上から明らかになるのは、数値データの場合にせよ文字資料のような定性的データの場合にせよ、縮約は、集約の作業を通してデータに含まれる情報量を「圧縮」するだけの手続きにはとどまらない、という点である。実際、要約統計量を算出したり文字テ

キストのデータからテーマを割り出したりしていく作業は、単に余分な情報を削ぎ落としていくプロセスなどではない。むしろ、分析者がデータに働きかけることによって、新たな構造や意味を創り上げる作業としての性格を持っているのである。

このような意味での縮約は、地図づくりの作業に喩えることができる。地図を作成する際には、三次元の地形を二次元の紙の上に投影することになる。そのようにして実際の地形を大幅に抽象化したかなりの部分を大胆に省略することによって、地図は、むしろ現実の地形や街区の構造あるいは鉄道の路線の配置などに関する、より効率的な理解を可能にする一種の「モデル」を提供しているのである（佐藤 2021: 202-203 参照）。同じように、要約統計量の算出やグラフの作成あるいはインタビュー記録を素材にしてテーマを割り出していく手続きは、現実の一面を鮮やかに描き出すことができる、明快な形で整理された「美しい」モデルを構築していくためには不可欠の作業なのである。²⁵

(2) 縮約と膨張——2つのベクトルの交錯

ところが、せっかく縮約の作業を通して数値データの分析結果をすっきりとした美しい形にしたはずなのに、その美しさや明快さを損なうような表示法になってしまう場合がある。その典型は、前稿で検討した、Excel で自動的に作成されるグラフに含まれる夾雑物であり、また本稿で見てきた、必要以上に桁数の多いコンマ以下の数値である。本稿ではこのように縮約とはいわば正反対のベクトルで作用してしまう手続きとその結果については、それを「データの膨張」と呼ぶことにしたい。²⁶

こうしてみると、これまで前稿と本稿で見てきた幾つかの事例（症例）は、上で言う本来の「縮約」とそれとは逆の性格を持つ「膨張」の2つのベクトルがそれぞれ独特の形で結合されており、それが結果として不可解な表現になっているのだと言えることができる。事実、それらの事例では、一見「データ自身に語らせる」あるいは「グラフ自身に語らせる」という理想を体現しているように見えるのだが、実際にはむしろ逆にデータの構造や意味が読み取りにくくなっているのである。

ここで改めてそれぞれの事例に見られる縮約と膨張のベクトルの概要を整理してみたのが、表6である（表の一行目には、縮約を視覚的に示す記号として下向きの矢印、膨張については上向きの矢印を添えておいた）。

25 しかし、各種の縮約やデータの加工には、当然、ここではかなり理想的に描き出した、現実把握のためのモデル構築の意図とは別次元で作用する政治的な思惑や力学が影響を及ぼしていることも多い（Starr 1983: 40, 51-52; Duncan 1984; Bulmer et al. 1991 参照）。

26 マイケル・スミスソンも data reduction の反意語として data enhancement を使用している。しかし、これは個々のデータの内容の詳細について検討していく作業（例えば、それぞれの対象者の回答の中味の検討）や主な数値データに他のデータを補足していく作業（例えば、自己評価に関する心理テストのスコアと自己評価に関する日記の記述を組み合わせる）などを指す用語であり、本稿における「膨張」とは基本的に性格が異なる（Smithson 2000: 21）。

表6 本稿と前稿における事例（症例）における縮約と膨張の交錯

	↓ 縮約	↑ 膨張（夾雑物） （無意味な「精度」・根拠薄弱な指標）
大学ランキング	・ 評点値 ↓ ・ 順位の数値	・ 小数点以下の評点値 ・ 明示化されない評価法
『ビジョナリー・カンパニー』 （「カルト的な企業文化」）	・ 評点の合計値 ・ 集計表での表示	・ 小数点以下の数値（.00）
『下流社会』 （「幸せを感じる時」）	・ 比率の値 ・ 集計表での表示	・ 小数点以下の数値 ・ 尺度化されない個々の指標 ・ 明示化されない項目の排列
『下流社会第2章』 （「心身の状態」）	・ 比率の値 ・ グラフ表示	・ 小数点以下の数値 ・ 尺度化されない個々の指標 ・ 明示化されない項目の配列 ・ 入り組んだ6本の折れ線

大学ランキングの場合は、幾つかの基準による評価をそれぞれの大学について1個の評点値に集約したという点では（その手続きの妥当性や信頼性はさておき）縮約の手続きが、型式上は採用されている。さらに、その評点を比較した上で単純な整数値という型式で表現された「順位」を示すという操作も縮約の作業だと言える。しかし、順位の基準として評点値を小数点以下1桁ないし2桁の数値まで使用しているのは、明らかに膨張である。また、中にはゴーマンレポートの場合のように評価法自体が公表されていない例もあるが、これも数値による縮約の効果を台無しにする膨張の手続きの一種だと言える。（就職力ランキングなどのように、評価基準が一応明示されているランキングの場合も、その基準に関する説明が不足していたり、基準自体の妥当性が疑われる例が多い。）

『ビジョナリー・カンパニー』の「カルト的な企業文化」の評価の事例については、3項目の評点を合算して1個の数値に集約したという点では、一応縮約の手続きが取られていると言える。また複数の企業文化の「カルト度」を一覧表（表1）としてまとめる手続きにも縮約としての性格がある。しかし、それらの縮約による「客観的で科学的な操作による評価」という印象は、全くと言ってよいほどに意味がない小数点以下2桁の数値（.00）が表に追加されることによって台無しになってしまっている。

また、本稿では『下流社会』に含まれていた、階層意識別の「幸せを感じる時」についての回答結果をまとめた表もとあげた。これも100名の回答者の質問票に対する回答を比率として集計した上で表（表2）にまとめたという点では縮約の手続きがとられている。しかし、例えば「上」の男性回答者について1名の回答を8.3パーセントとするという例に見られるように不必要な小数点以下の数値を含む比率の数値が使用されている。また、この表では、個々の選択肢が尺度化という集約の手続きも経ずにそのまま指標として使われている。さらにそれらの項目の配列の仕方も明示されていない。これ

らの3点は明らかに膨張の作用をもたらしており、また示された数値データはほとんど無意味なものになっている。

なお前稿では、『下流社会第2章』に含まれていた不可解な折れ線グラフの事例を取りあげた。その折れ線グラフは、6種類の就労形態の男性回答者の心神状態に関する質問への回答を集計した結果を示したものだと言う。本稿の表2と同じように、この折れ線グラフについても、集計結果としての比率の数値をグラフ形式で表示しているという点に限って言えば縮約の手続きが一応は取られている。しかし一方で、このグラフの場合は、本稿の表2と全く同様に、小数点以下の数値、尺度化されない個々の指標、明示的ではない項目の配列という3点に加えて、6本の折れ線が入り組んで込み入った図解表現となっている。つまり、膨張の手続きがいわば「てんこ盛り」になることによって「超絶技巧」とも言えるほどに不明瞭で不適切な図解表現になっているのである。

(3) 膨張の2つの背景——リテラシー不足と目くらまし

前稿ではナフリックの6つの教訓の1つとして「デザイナー思考を心がける」というものをあげた。表5でとりあげた事例の文献の著者たちは、いずれの場合も、デザイナーとしては明らかに失格である。というのも、彼らは、本来、縮約の手続きを通してスッキリとした明快なモデルの構築を心がけるべきであった。しかし実際には、不必要な夾雑物をグラフや分析の結果の上に「ゴテゴテ」と貼りつけて膨張させることによって、決して美しいとは言えない装飾過多のモデルを作り上げてしまっているのである。

このような膨張の背景には、幾つかの事情や理由ないし思惑があると考えられる。

1つは、基本的なりサーチ・リテラシーの不足ないし欠如である。前稿で指摘したように、リサーチ・リテラシーは単に書き手の側の読み書き能力だけを意味するわけではない。それに加えて、読み手の立場にも十分に配慮した双方向的なコミュニケーションをおこなうための読解能力が必要になってくるのである。前稿で例示した不可解なグラフや本稿で見てきた不必要な小数点以下の数値による表示は、見方によってはいずれも一種の「うっかりミス」とでも呼べるものであろう。その中には、微修正ないし「ひと工夫」という程度で十分に修正が可能なものも含まれている。しかし、そもそもその種の「ミス」の存在に気がつかないという点では、基本的な素養としてのリサーチ・リテラシーが欠如していると言わざるを得ないのである。

一方で、データの膨張の背後には、このような「うっかりミス」の範囲を越えた、さらに根が深く、きわめて深刻な事情が存在している場合も多い。その点は特に、膨張の操作が確信犯的に一種の「目くらまし」の一環として加えられているとしか思えないケースについて指摘できる。そして、その「目くらまし」には、少なくとも2種類のものがあると考えられる。また、それら2つのタイプの目くらましが併用される場合もあ

る。1つめのタイプは、膨張が調査結果の科学性や客観性に関する演出効果を意図して採用されている場合である。それに対して2つめのタイプは、分析結果の内容の乏しさを覆い隠した上で「やっている感」を醸し出すためにあえて読み取りにくい図解表現やコンマ以下の数値が使われる、というケースである。

(4) 厳密科学に期待される明快さと「厳密性」

1つめの科学性の演出効果という点に関しては、自然科学系の(つまり「本物」の)厳密科学それ自体に、明快なモデルの構築に加えて精妙な測定結果の提示という2つの側面があるという点について理解しておく必要があるだろう。

物理学をはじめとする「厳密科学」の発見事実にはしばしば見られる1つの特徴は、数値をはじめとする各種の経験的データにもとづいて作成される単純明快で「美しい」モデルや法則である(クリース 2006²⁷)。その典型例の1つには、DNAの二重らせんモデルがある。また、ランドバークが*Foundations of Sociology*で何度となく言及したアインシュタインによって提示された特殊相対性理論から導き出された $E=mc^2$ という、質量とエネルギーの関係を示す等式も、科学研究の1つの到達点としての美しく簡明なモデルを象徴するものだと言えるだろう。さらに、科学研究の実態とはかけ離れた形で、あたかもステップ・バイ・ステップ方式であるかのように理想化・理念化された仮説演繹法的手順などもこれに含めて考えることができるかも知れない(Medawar 1964, 1967)。

また、その科学研究の成果として提示される各種のグラフも、数値データから割り出される規則性や傾向を明快に示しているという点で、科学的探究やそれによる数値データの縮約の力を見事に示す図解表現であると言える。さらに卑近なレベルでは、社会現象に関わる情報に関して、文字テキストなどのような入り組んだデータを「きたないデータ」と見なすのに対して数値データを「きれいなデータ」、またモデル通りの分析結果について「きれいな結果」などと呼ぶ慣習も、厳密科学からの連想によるものであろう。

もっともその一方で、これまで本論考で何度か見てきたように、厳密科学のもう1つの特徴は、精度の高い客観的な測定と分析である。それがまた、厳密科学(exact science)という名称の由来でもある精密な測定(exact measurement)とも密接に関連している。

この場合は逆に、大胆な省略による簡明なモデルの構築というよりは、むしろ可能な限り詳細な情報の取得と提示に価値があるのだとされる。実際、自然科学の世界では、より正確な測定精度を求めて、測定原器については100年以上にわたって何度となく改訂がおこなわれてきた。また、2000年代に入ってからほとんどの測定単位が物理定

27 時にはそれによって科学者が自然界に本来存在する乱雑なパターンを見逃すことにもつながってきたとも言われている(シュービン 2021: 215)。

数や原子の性質によって再定義されるところまで進んでいる。そのようにして達成された世界共通の基準によって精密な測定が可能であることが前提とされている (クリース 2014)。また、そのような測定精度の進化は、科学研究がさまざまな自然の「秘密」を解き明かし、人間の五感ではとうてい捉えきれない現象をその細部にわたって描き出してきたと想定されているのである。

自然科学系の研究では上で述べたような縮約の手続きによる簡明な描写と分析とナノ単位にいたるまでの精密な測定という2つのベクトルに調和が取れていることも多いだろう。実際、自然科学におけるミリ単位ないしナノ単位の測定精度は、まさに厳密科学こそがなし得ることである。それに対して、社会現象を対象にする研究においては、単にそのような精度による測定は到底なし得ないだけでなく、2つのベクトルは、とすればアンバランスなものになりがちなのである。

(5) 「やっている感」の演出と読み取りにくさ——安易な縮約と道具優先主義

それもあって、科学研究の形式を模倣して見かけだけを取り繕うという事態も生じてきた。その場合に示される図解は考え抜かれた末に出来上がってくる美しい理論モデルなどではない。また、一見精度の高い数値も、不必要な「コンマ以下」の情報をノイズ (夾雑物) として含むコケオドシでしかない。つまり、数値データや図解表現は時として見かけ上の科学性と客観性、言葉を換えて言えば「意味ありげ」な科学性の印象を醸し出す上で利用されてきたのである。こうして、縮約の作業は意味を刻み出すための手続きではなく、しばしば、むしろ意味ありげなグラフや小数点以下の数値などの膨張の要素を含む解析結果を生み出すことになる。

これらの事実の背景には、データ分析の本質的な目的に関わる問題が含まれている場合も多い。つまり、数字やグラフで語ろうとする際に、社会現象の姿を浮き彫りにするという目的とはかなり性格が異なる、他の動機が介在している例が少なくないのである。

それは、例えば「(本格的な調査を) やっている感」の演出であつたり、研究者の場合には研究業績の量産であつたりする。研究活動を、学術的知識の向上を目的としておこなう活動としてではなく、もっぱら研究成果によって安定したポストの確保やキャリアアップを目指すための手段として考えるような傾向を指して手段優先主義 (instrumentalism) と呼ぶことがある。また、長期的な取り組みが必要となる研究テーマを忌避し、手っ取り早く論文にまとめられるようなテーマを選択する傾向は、短期決戦主義 (short-termism) と呼ばれる (クリメッキ 2018: 327-331)。そのためには、調査結果を効率的に伝えるよりは、むしろあえて読み取りにくい図解表現を使ったり小数点以下の数値を駆使した方が「目くらし」としては効果的であるかも知れない。

不可解なグラフや不必要な「精度」の背景には、そのような、読み手の側の立場や視点を軽視して、むしろ書き手の側の都合を優先するような手段優先主義と短期決戦主義があると考えられるのである。

IV 「こたつ論文」からの脱却 (処方 : How to)

当然ではあるが、本稿と前稿で指摘してきたさまざまな問題は、必ずしも単に図解による表現の仕方を工夫したり、数値の「精度」の表示法について見直すという程度の対応で解決できるような話ではない。不可解なグラフ表現や小数点以下の数値に対する強迫的とも言える依存傾向の背後には、そのようなテクニカルレベルの対応の範囲をはるかに越えて、より本質的な、各種の数値データおよびそれらのデータが示す社会現象に対する基本的な向き合い方ないし「つき合い方」に関わる深刻な問題が潜んでいるのである。

社会現象の数値ないし数値化に対する信頼感は長年の歴史を経て形成されてきたものである (Starr 1983; Porter 1986; 1995; Desrosières 1998; レイ 2020²⁸)。したがって、その信頼感の歪んだ派生物とも言える「数値信仰」それ自体をより健全な方向に修正していくことは、明らかに小論の範囲を越える。ここでは、本論考で取りあげてきた数値データの縮約の手続きと数値データそれ自体の性格という2つの問題に焦点を絞った上で、数値データを元にしたより効果的なコミュニケーションを図る上で不可欠となる対応策についてかなり控えめの提案をしてみたい。

1. 数字とグラフで語ることの意義について再認識する

(1) 数字が語る 対 数字で語る

先に「データそれ自身に語らせる」ないし「グラフそれ自体に語らせる」という考え方の落とし穴について論じた際にも強調したように、数値データやその図解表現は決して現実世界を曇りなく映し出す鏡のようなものではない。また、データそのものは決してみずから語ることなどない。人間が特定の枠組みにもとづいて適切な解釈をし、ふさわしいレトリックを使うことによってはじめて「データそのものが語る」ように見せかけることができるか、あるいはまた「データとともに語る」ことができるだけなのである。「データ自身が語っている」ように見えるのは、知らず識らずのあいだに身につけてしまった特定の物の見方に囚われているために、むしろ「データに欺かれた」(データ

28 ランドバークに象徴される厳密社会科学への憧れも、そのような数値への信頼および社会現象に関する統計的表現の普及を抜きにしては考えられないだろう。なお、数値自体に対する信頼と数値の使用法(例えば、国勢調査の統計データの徴税や徴兵のための使用)に対する信頼は全く別次元の問題である。

に「騙られた」時であるとさえ言える。

言葉を換えて言えば、我々が目指すべきなのは、数字がみずから語ることを受け身的に待ち受けるのではなく、むしろ可能な限り明快にかつ正確に「数字で語る」ことなのである。実際、言葉で語る際にさまざまな工夫が必要であるように、数字で語る際にも、さまざまな工夫とそれ相応のトレーニングが必要なはずである。

同様の点は、グラフによる数値データの視覚化という手続きについても指摘できるだろう。我々は数値データの集まりが図解形式で表現されたものを目にすると、ともすれば、人間の五感では捉えきれないようなデータの構造が見事に解き明かされ、またそれによって現実社会のあり方が正確かつ忠実に描き出されていると思ひ込みがちである。言葉を換えて言え場、そのようなグラフを目にした時には、「グラフが語って」いるように思えてくるのである。

しかしながら、前稿で、明らかに不適切な事例も含めてさまざまな円グラフ、折れ線グラフ、帯グラフの表現方法について見たように、グラフそれ自体が語ることは決してない。作成者の側（見せる側）が、コミュニケーション・ツールとしてのグラフの意義について十分に認識し、また特定の読み手（見る側）の立場と視点に立ってさまざまな技法を駆使することによって、初めて効果的に「グラフで語る」ことができるのである。また、その時にこそ、グラフを見る側の目には「グラフが語って」いるように思えてくるのである。

その意味では、データの「可視化」という言い方は必ずしも正しくない。というのも、グラフをはじめとする図解表現で我々が目指しているのは、既に確固としたものとして存在している数値データ、またそれが言及している社会的現実の構造や意味を単に「見える化」とするというわけではないからである。むしろ、我々の側が数値データに対して積極的に働きかけることによって、その構造や意味を創り上げているのである。

(2) 造形作業としての縮約

その意味で「数字で語る」手続きには、芸術家とその素材である木材、岩石あるいは金属から作品を彫り出していく作業によく似た面があると言えるだろう。

例えば、統計調査における基本的な手続きとしてのデータの縮約は、必ずしも、データの堆積の中からひたすら余分な要素を削ぎ落として「真実の構造」や「真のパターン」ないし唯一正しい意味を浮き彫りにしていくプロセスなどではない。むしろ、素材に対して作成者が独自の視点で向き合うことによって新たな作品を創り出していく彫刻の作業に似た側面があると言える。あるいはまた、宝石彫刻師が加工や研磨の作業を通して原石の中から特定のパターンを持った宝石細工を削り出していく作業に似た面があるとも言える。実際、前稿でナフリックの6つの教訓の中に「デザイナー思考を心がけ

る」というものがあつたように、データ分析には、本来きわめて創造的で能動的なプロセスが含まれているのである。

前稿と本稿で見てきた、不可解なグラフや測定精度過信症的な傾向の背景には、このような創造的なプロセスが見落とされていることによって、統計処理における「基本の基本」が軽視されているという事情がある場合も多いだろう。つまり、これらの場合は、何らかのパターンや意味の判別ができるような形でデータを加工していくという作業の本質が見落とされているのである。

要するに、本来はそのままでは混沌としたもののように見えるデータの山（堆積）から何らかのパターンや「意味」を丹念に抽出していく際の第一歩であるはずのデータ縮約のプロセスが、むしろ逆に「意味ありげ」な、まさにあたかも「数字自身が語っている」ようなイメージを醸し出すために使われてしまうのである。だからこそ、縮約とは正反対の「膨張」のベクトルが働いてしまうのだとも言えよう。

2. データの作られ方とその文脈について理解する

(1) 数値への「還元」と数値の「縮約」

以上の点は、先にふれた「データの縮約」という手続きについて新たな視点から考え直していく必要性をも示唆している。縮約の原語は **reduction** であるが、この言葉には還元主義 (reductionism) を連想させてしまう可能性があり、やや紛らわしいところがある。一般に還元主義というのは、高次の現象をより低次の現象によって説明しようとする試みのことである。また複雑な現象について、それを構成する基本的な要素の単純な原理で説明しようとするアプローチを指す場合も多い。典型的な例としては、全ての科学は原理的に物理学の法則によって説明可能であると主張したり、組織や集団に関わる現象の全ては個人の行動によって説明できると主張するような発想（いわゆる「方法論的個人主義」）などがあげられる (Giddens 1979)。

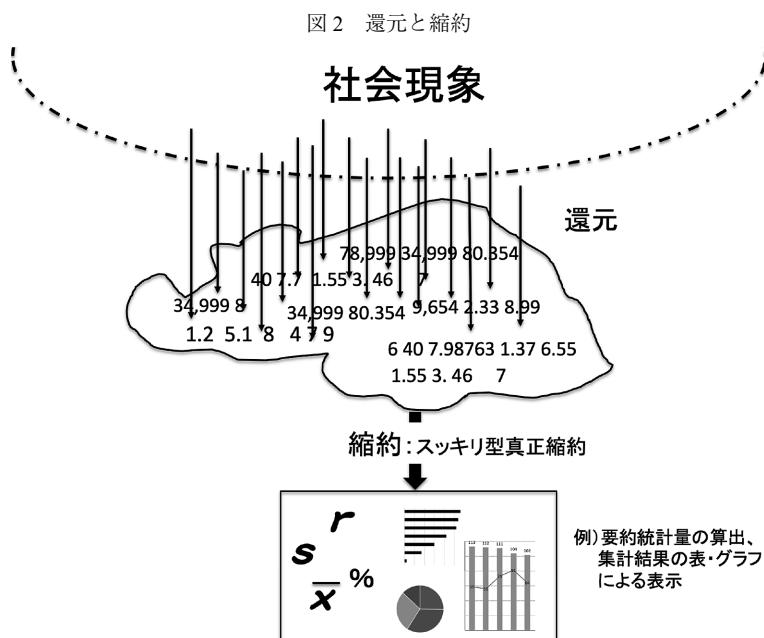
ある意味では縮約の手続きの前提となる、社会現象に関わる対象の数値化自体が、そのような意味での **reduction** つまり還元の手続きを含んでいるとも言える。

実際、さまざまな社会現象を数値化する作業というのは、その対象となる組織や個人が持っている多様な側面を数字という抽象的な次元で処理できる部分に着目して単純化して変換（還元）するという手続きが含まれている。したがって、そのような形で数値化できない部分については大幅に省略していくことになる。統計的調査は、その還元の手続きを通して記述や説明の一般化を目指すところに本質的な意義がある場合が多い。また、先に地図の喩えを使って解説したように、そのような大胆な抽象化と省略の作業を通して、個々の事例の詳しい情報だけに目をとらわれている限りは見えてこない全体的な傾向や一般的なパターンを浮き彫りにしていくこともできる。本稿で見てきた縮

約, つまり要約統計量を算出したり, 数値データをグラフや表の型式で整理して表示することは, まさにそのような還元の手続きを通して「ビッグ・ピクチャー (全体像・大局観)」を明らかにしていく作業であると言える (Ragin 1994: 92; Smithson 2000: 22)。

以上のような点からすれば, 数値データを分析していく作業について理解する際には, 2つの意味での reduction を区別しておく必要があると言えるだろう。つまり, 統計的調査には多くの場合, 数値への「還元」と, 数値の縮約, つまり既に数値化されたデータの集積の中からその構造や意味を明らかにしていく意味での縮約という2種類の作業が含まれているのである。

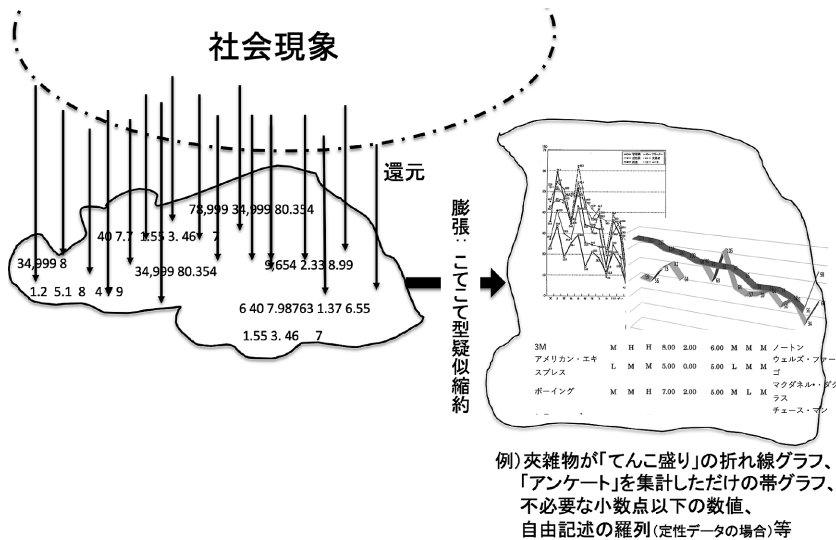
その2つの作業, つまり「数値への還元」と「数値の縮約」という手続きの関係を図解してみたのが, 図2である。



先に指摘したように, 数値の縮約の作業には単なる単純化や情報の圧縮というだけでなく, むしろ新しい構造や意味を創り出していく, あるいは彫刻のように「彫り出して」いく側面が含まれている。つまり, 縮約には, そのままでは単なる無秩序な数値の堆積にしか見えない場合も多いデータの集合に対して新しい構造的な秩序と意味を付与していく「拡張」とでも呼ぶべきプロセスが必然的に含まれているのである。(当然ではあるが, 縮約以前の段階でおこなわれる還元もまた, 単に既存の基準を適用してカウントするというような単純な手続きではない。調査対象を特定のカテゴリー (例えば, 人種, 所得階層, 性別等) に分類し, また指標値・尺度値を割り当てたり重み付けしたりするという点では, むしろきわめて創造的 (時には作弄的) なプロセスが含まれているのである (Starr 1983: 29-49)。

一方で、本稿で解説してきた「膨張」は、一見縮約と同じように要約統計量や図解表現を使ってはいるものの、実際には、新たな意味を創造するというよりは、むしろ「精度の高い測定や判定・評価」という意味ありげな印象を演出しているに過ぎない。その意味で、「膨張」は本来の縮約とは似て非なる、いわば擬似的な縮約の手続ききだと言える。この事情を図解してみたのが、図3である。

図3 膨張=擬似的縮約



出所：一部は三浦（2007）およびコリンズ＆ボラス（1995）

この図に示したように、膨張の場合も、本来の縮約と同じように、グラフなどの図解表現を用い、また分析や評価の結果と称してコンマ以下を含む特定の数値を示したりはしている。しかし、その情報の集約の程度は、数値形式に還元された段階のデータのレベルとあまり変わるところがない。実際、そのような擬似的縮約では、むしろ余計な「ノイズ」のような夾雑物が加わることによって、社会現象の構造や意味あるいは「ビッグ・ピクチャー」は一向に見えてこない。したがって、図2および図3にも示したように、本来の縮約は「スッキリ型真正縮約」、膨張については「こてこて型疑似縮約」とでも名づけて区別しておいた方が良いのかも知れない。

なお、本稿では主に数値データの分析を中心にして縮約や膨張の性格について解説してきたが、実際には文字テキストなどの定性データについても、同様の点が指摘できる。先に定性的データの縮約の例としては、コーディングなどを経てテーマを見出した複数のテーマ間の関係を割り出していく作業をあげた。これは、本来の意味での縮約の手続ききであると言える。それに対して、例えば、「アンケート」などの自由記述欄の項目を単に羅列して済ませるといような「分析法」もある。それは、ここでいう膨張、すなわち擬似的な縮約に過ぎない。実際、その「ベタな記述」の羅列は一見意味あ

りげではあっても、そこからは社会現象の意味や構造は一向に浮かびあがってはこない。単に「やっている感」を演出するためのギミックとしては効果的ではあるかも知れないが、とうてい縮約ないし「分析」と呼ぶには値しないのである (佐藤 2021: 117-119 参照)。

(2) 還元主義の罠——「粗悪データ身につかず (easy come, easy go)」

データの「作られ方」を念頭において分析を進めていく際には、縮約の意義について理解することに加えて、数値データへの還元にもなう落とし穴についても理解しておく必要がある。また、数値データの入手先や入手方法によって還元主義的な度合が異なる点にも注意が必要である。

先に述べたように、数値への還元は社会現象の「ビッグ・ピクチャー」を把握し描き出す上では非常に効果的であることが多い。しかし、その一方で、数値化やその後に続く数値データの縮約の手続きは、個々の事例の独自性と文脈を削ぎ落としてしまい、生きている個人や社会・集団を単なる1つあるいは複数の数字に文字通り「還元」してしまう可能性がある (Smithson 2000: 22)²⁹。そのような場合、社会現象からは微妙な雰囲気も色も匂いも消え失せ、無機質で無味無臭、手触り感もほとんどないものになってしまいかねない。また、その種の安易な還元操作によって作り出された数値データのみで描き出される個人、集団あるいは組織は、その姿や面影どころか、名前も顔も見えない、血の気が通っていない、青ざめた生ける屍 (ゾンビ) のように見えることさえある。

数値データを利用する際には、そのような還元主義的な傾向に対して十分に注意を払う必要がある。その傾向は、現在のように、インターネットを介して電子化され、多くの場合にはあらかじめ相当程度の加工がなされてしまっているデータが大量に入手できるような状況では、さらに加速されてしまう可能性がある。というのも、数値データそれ自体の「つくれ方」の経緯や文脈が抜け落ちてしまう場合も少なくないからである。³⁰

その種の疑わしい出所を持つデータの例としては、例えば幾つかの国で常習的になされてきた GDP 偽装 (コイル 2015; カラベル 2017) や日本における多くの府省による「統計不正」 (佐藤 2021: 300-301) があげられるだろう。このようなデータの中には、時に、各種の

29 もちろん、全ての数値化がそのような還元主義的な結果に終わるわけではない。例えば、病理検査が各種の指標を駆使して症状の根底にある病根を立体的に描き出すことがあるように、数値によってむしろ複雑な姿が浮かびあがってくることもある。これについては、Smithson のデータ拡張 (data enhancement) の議論を参照 (Smithson 2000: 21-22; Ragin 1994: 92 をも参照)。

30 社会学者のジュリアス・ロス (Julius R. Roth) は、1966 年にデータ収集を請負任せにすることががともすればデータの質という点で致命的な問題を引き起こすことがある点について警告を発していた (Roth 1966)。50 年以上の歳月を経て今や問題はそのような範囲を越えている。実際、ロスが想定していたのは、例えば学生や若手の教員にデータ収集を請け負わせるような事態であった。その場合、少なくとも請け負った調査担当者の顔は見えているのである。一方、現在ウェブ上から入手できる数値データは、調査を請け負った者の顔どころか、その出所、素性さえ明らかではないことが多い。

経済分析あるいは EBPM の提唱者たちの主張を空虚なものにしてしまうものが含まれている。

この点に関連して、2018 年 12 月から次々に明らかになった政府の基幹統計に関する「統計不正」に関連して、神林龍・一橋大学教授は、次のように語っている——「EBPM の掛け声とは裏腹に、政策立案者も政策遂行者も研究者も、統計作成を軽視するようになってきたとすれば、問題はかなり根深い」(神林 2019)。そして、その統計不正を受けてさまざまな改革がなされてきたとされるにもかかわらず、最初の不正事件が発覚したちょうど 3 年後の 2021 年には、新たに国土交通省の「建設工事受注動態統計」に関して、原票の記載の一部を消しゴムで消して改竄するなどという、かなり稚拙かつ悪質な統計不正が発覚したのであった。

先に、インタビュー記録などの定性的データと比較して数値データには「きれいなデータ」という印象がつきものであると指摘した。しかし、インターネットを介して容易に入手できる数値データの中には、一種の「データ・ロンダリング」の手続きを通して「きれいなデータ」のように偽装されている例が少なからずあると推測できる(国交省のウェブサイトでは、建設工事受注動態統計については、2019 年 4 月分から速報の公表が取り止められているのだが、それまでは広く公開されていたのである)。

マネー・ロンダリング(資金洗浄)では、違法な手段を通して得られた資金に操作を加えて一見「きれいなお金」に加工してしまう。それと同じように、ここで言うデータ・ロンダリングの場合は、その根拠も「素性」も明らかではないデータが数値形式をとっているというその理由だけで、いつの間にか「きれいなデータ」として通用してしまう³¹。

当然ながら、これは「ウェブ由来のデータが全て信用に値しない粗悪な情報でしかない」ということを意味しない。ここで強調しておきたいのは、二次的(ないし三次以降の n 次的)な情報を利用する際には、その出所やデータの質そしてまた制限条件も含めて的確に判断することができる「データ鑑識眼」を養っていく必要がある、という点なのである(佐藤 2021: 293-319)。実際、さもないと、どのような高度な解析技法を用いたとしてもそのようなジャンクデータからは GIGO(ガーベージ・イン・ガーベージ・アウト: 屑入れ屑出し)的な分析結果しか得られないだろう。諺に「悪銭身につかず(easy come, easy go)」と言うものがある。それをもじって言えば次のようになるだろう——「粗悪データ身につかず」。

(3) 厳密社会科学と「こたつ論文」のあいだ

31 英語でいう data laundering は、通常、盗用されたデータの出所を明示せずに利用したり売買する行為を指す。

ランドバーグがかつて夢見ていた厳密社会科学の理想を一方に置き、他方に前稿と本稿で解説してきた不可解なグラフや「コンマ以下」の数値に対する執着を置いて見た場合、両者のあいだには絶望的なほどの遠大な距離があることは明らかであろう。ランドバーグ自身が最も雄弁な論客の1人であった社会調査の「科学化」と定量化は、一面では各種の統計分析の充実をもたらしたものの、他方では、かなり疑わしい調査法を操る「アンケート産業」を生み出し、また数値信仰を蔓延させる上でも少なからぬ役割を果たしてきた(川喜田2006: 171-174)。それがまた、極端な場合には各種のデータ・ロンダリングを発生させることもあったと言えるだろう。

前稿でもふれたように、近年になってその傾向は、Excelをはじめとする便利な表計算ソフトウェアや Google Forms などの各種のオンラインサービス、そしてまたインターネットを介して瞬時にアクセスできるようになった電子的データベースの拡充によってますます加速されているように思えてならない。そして、ある種の論文であればそれらの情報を活用して「プッシュボタン式の統計調査」によって効率的に量産していくこともそれほど難しくはない。

それらの論文の中には、いわゆるこたつ記事を連想させるものが少なからず含まれている。「こたつ記事」というのは10年ほど前から主にジャーナリズムの世界で使われてきた言葉であり、それを造語した本田雅一によれば、「ブログや海外記事、掲示板、他人が書いた記事などを“総合評論”し、こたつの上だけで完結できる記事」のことだという(本多2021)。つまり、苦勞して物事が起きている現地に取材に赴かなくても手元にPCとインターネット環境さえあれば、それこそ「こたつ」の上でも書けてしまうような記事を指す。

同じようなことが、社会調査についても指摘できる場合がある。実際、今や社会現象が生じている現場に直接出かけて行かなくても、手元のPCあるいはスマートフォンを使えば、各種の数値情報は言うに及ばず、膨大な量のテキストデータや画像情報も入手できる。また、各種の有料のデータベースサービスも充実している。

自分で調査をおこなう場合であっても、インターネットを介して相手側から情報を提供してもらうことができる。例えば、新聞社ないし出版社系の大手メディアであれば、回答に応じないことの「たたり」(川喜田2006: 172)を恐れている大学に依頼してウェブを介して調査票に記入してもらうことで大学ランキングを作成することも容易であろう。かくして、「こたつ調査」による「こたつランキング」が出来上がってくわけである³²。また、Google Formsのようなプッシュボタン式の「アンケート」調査と集計のためのサービスをうまく使えば、「こたつ調査」によって「こたつ記事」式のアンケート論文

32 学長、学部長、就職部長等の大学幹部、学生であれば数名にインタビューをする程度である種の現場感覚の味付けをすることも可能だろう。

(こたつ論文)を量産することもそんなに難しくはないだろう。

そのような便利な時代に、わざわざ時間と手間と労力をかけ、また時には対象者とのやりとりでジリジリするような思いをしながらインタビューをおこなったり、質の高いアーカイブ的な統計データを入手するために調査現場を訪れることは徒労としか思えないかも知れない。何よりも、そのような「手間暇をかけた」調査では業績点数を効率的に稼ぐことは難しくなってしまうのである。

(4)「こたつ」の上のスマホと PC を捨てよ、町へ出よう

シカゴ大学を拠点にして活躍した「シカゴ学派」と呼ばれる社会学者たちのグループは、1920年代から30年代にかけて数々の優れた都市民族誌(都市社会を対象にしたフィールドワークの報告書)を生み出した。その社会学者たちを育てたロバート・パークは、ある授業で学生たちに向かって次のように述べている。

君たちは、これまで何度かこんなことを言われてきただろう——「図書館に行って本を漁って山のようなメモをとり、体全体に埃をつけてくるように」。自分で研究テーマを見つける時には、ちゃんとした記録があればどこからでも選んでいい、とも言われてきただろう。たとえそれが、疲れはてた役人が作った面白くもなんともない質問表に何かの補助が欲しい人がしぶしぶと答えたような、そんな記録があるところとか、小うるさい一人よがりの社会改良家とかやる気のない事務屋が記入したようなかび臭いおきまりの記録の山だとしてもね。そんなことが、「実際の調査で手を汚すこと」って呼ばれているんだ。……でも、もう1つどうしても必要なものがあるんだ。自分の目で見ることだよ。一流ホテルに出掛けていってラウンジに腰掛けてみなさい。安宿のあがり口に腰をおろしてみなさい。ゴールドコーストの長椅子やスラムのベッドに腰をおろしてみなさい。オーケストラホールやスター・アンド・ガーター劇場の座席に座ってごらんない。

要するに、諸君、街に出ていって諸君のズボンの尻を「実際の」そして「本当の」調査で汚してみなさい (MaKinney 1966: 71)。

このパークのアドバイス(ないし「檄」)では、都市の現場でおこなうフィールドワークが「『本当』の調査 (real research: イタリックは原著)」であるとされ、それと図書館や記録が保管されている場所でおこなう調査が対比されている。しかし、1920年代のパークの、シカゴ大学の教室における発言から約100年の歳月を経た今では、パークが言う「体全体に埃をつけて」おこなうような文献・資料調査を実施することさえ稀になってきている。³³

33 この点については、1979年に米国学術団体評議会 (American Council of Learned Societies) が発表した報告書である *Scholarly Communication* の中で描かれている「未来の研究室」のイメージが興味深い。(American Council of Learned Societies 1979)。この冊子は学術ジャーナルの増加や購入価格の高騰などにともなって、大学図書館の書籍購入が危機的な状況に陥っていったことなどを背景にしておこなわれた調査にもとづいて編まれた報告書である。同協議会は主として人文・社会科学系の団体が参加していたので、書籍刊行の危機は特に切実な問題だったのである。同報告書では、必要な文献を「ブラウゾ

先にあげた本田雅一は、次のようにも述べている——「ぬくぬくと暖かい部屋の中で勞せず記事を書いている、とイメージしたことも否定しない。だから“こたつ”記事なのだ」。そのこたつ記事と同じように、インターネットを介して容易に入手できる情報だけを使っている限りは、数値データは、社会現象を曇り無き映す鏡であるどころか、ますます調査する側が見たいものだけ見るような歪んだレンズのようなものになってしまうことはないだろうか。

先に強調したように、もし良質の数値データを得ようとするならば、そのデータの作られ方や情報が数値に還元される以前の社会生活の文脈にまで遡って理解しておく必要がある場合が少なくない。そのためには、居心地の良い「こたつ」の布団から身を引き剥がして、外の冷たい風にあたってみることがどうしても必要になってくるに違いない。

要するに、学生諸君、(新型コロナウイルス感染症収束の暁には) スマホと PC を捨てて町へ出てみようではないか！

【補論】データ縮約の「べし・べからず集」(Do's and Don'ts)

前稿と本稿では、プッシュボタン式統計調査の際に初学者が陥ってしまいがちな幾つかの落とし穴について解説してきた。あわせて、より効果的なデータ縮約の方法についても指摘してきた。本論考を締めくくるにあたって、以下では、それらを一種の「べし・べからず集」(Do's and Don'ts) 的にチェックリスト形式でまとめてみた。学生諸氏は、学位論文を提出する前などに、以下のチェックリストを使用してみると良いだろう。

なお、「べし (Do's)」については、チェック欄の□の後に○を、「べからず (Don'ts)」については×を付して区別しておいた。

1. 図解表現による縮約の「べし集」：総論

以下は、前稿の記載の一部を再録したものである (佐藤 2022: 136)。これは、Knaflic (2015) が数値情報を図解によって効果的に表現していく際のポイントを6つの教訓として整理したものである。

ㄨ 管」上で見たり、コンピュータに接続された「タイプライター」でプリントアウトして読むことが想定されている。また、書籍については、それぞれの図書館に「電話」すれば「マイクロフィッシュ」の形式で保存しているデータを元にして印刷するというようなサービスも描かれている。それらの電子化の進展によって、研究者は、図書館に赴いて本を探したり、探していた本が見つからずに数週間またされたりすることなくなることによって、「考えたり書いたり、そして教育にもっと時間が使えるようになるだろう」としている (American Council of Learned Societies 1979: 31-32)。さて、果たして、この報告書に描かれていたように、電子化のお陰で我々は今日じっくりと考える時間を持つことができているのであろうか。

- ☐ (○) コンテキストについて理解している
- ☐ (○) 最も適切だと思われる視覚表現を選択して採用している
- ☐ (○) 夾雑物を排除している
- ☐ (○) 読者に注目してもらいたい部分を強調している
- ☐ (○) デザイナー思考を心がけている
- ☐ (○) ストーリーを明確に語っている

2. 図解表現による縮約の「べし・べからず集」：各論

(1) 円グラフ

- ☐ (×) 原則として使わない (どうしても必要な場合を除く)
- ☐ (○) 棒グラフ等で代替できる場合は代替する
- ☐ (×) 3D 表示にはしない
- ☐ (×) Excel 等のデフォルト設定で使用しない (夾雑物を排除する)
- ☐ (○) 「パイ」の配置や順番が一定の規則ないし慣行に従っている
- ☐ (○) 読者に対して伝えたい内容と「パイ」の配置や順番とのあいだに整合性がある

(2) 折れ線グラフ

- ☐ (×) 必要性・必然性が無い限りは原則として使わない
- ☐ (○) 棒グラフ等で代替できる場合は代替する
- ☐ (○) 一連の数値で代替できる場合は代替する
- ☐ (×) 3D 表示にはしない
- ☐ (×) Excel 等のデフォルト設定で使用しない (夾雑物を排除する)
- ☐ (○) 横軸項目の配置や順番が一定の規則ないし慣行に従っている
- ☐ (○) 読者に対して伝えたい内容と横軸項目の配置や順番とのあいだに整合性がある

(3) 帯グラフ

- ☐ (×) 「アンケート」等の一次的な集計結果の表示には使わない
- ☐ (○) 横棒ないし縦棒のカテゴリの配置や順番が一定の規則ないし慣行に従っている
- ☐ (×) Excel 等のデフォルト設定で使用しない (夾雑物を排除する)

3. 小数点表示ないし社会的測定の精度に関する「べし・べからず集」

- ☐ (×) 無意味な「精度」を強調しない
- ☐ (○) 有効数字に関する慣行ないし「相場」について理解する
- ☐ (×) Excel 等の出力結果を「コピペ」しない
- ☐ (○) 物理的測定と社会的測定の本質的な違いについて理解する

- ☐ (○) 尺度の測定レベル (佐藤 2021: 313-319) について理解した上で数値表現を使用する
- ☐ (×) 「統計的に有意な差」を客観的で科学的な指標などとは考えない
- ☐ (○) p 値によって統計的有性を示す場合は、単に「変数間に何らかの関係がある」というだけでなく、その関係の程度 (効果量等)、方向性、形式、関係に対して影響を与える条件について解説する (Edwards, J., and Berry, J. 2010; トゥーリッシュ 2022: 218-219)

*本稿の作成にあたっては、以下の研究助成を受けている——JSPS 科学研究費補助金 (課題番号 19K02144)

引用文献

※ウェブ上の情報は全て 2022 年 2 月 26 日にアクセス

- 石川真由美編 (2016) 『世界大学ランキングと知の序列化——大学評価と国際競争を問う』京都大学学術出版会.
- 岩井八郎 (2005) 「開き直る『中の下』を浮き彫りに」(『下流社会』書評) 『日本経済新聞』2005 年 10 月 23 日付.
- カラベル, Z. (北川知子訳) 『経済指標のウソ——世界を動かす数字のデタラメな真実』ダイヤモンド社.
- 川喜田喬 (2006) 「調査屋の心構え——産業社会学とフィールド調査」小池和男・洞口治夫編 『経営学のフィールド・リサーチ』日本経済新聞社.
- 神林龍 (2019) 「毎勤統計不適切調査の背景 政策立案と遂行の分化映す」 『日本経済新聞』2019 年 1 月 28 日付.
- クカーツ, U. (佐藤郁哉訳) (2018) 『質的テキスト分析法』新曜社.
- クリース, R. P. (青木薫訳) (2006) 『世界でもっとも美しい 10 の科学実験』日経 BP.
- (吉田三知世訳) (2014) 『世界でもっとも正確な長さや重さの物語』日経 BP.
- クリメッキ, R. (2018) 「個人的体験としての REF」佐藤郁哉編著 『50 年目の「大学解体」20 年後の大学再生』京都大学学術出版会, Pp.307-337.
- 経済法令研究会 (2012) 『モンドセレクションガイドブック食品編——モンドセレクション 50 周年記念』経済法令研究会.
- コイル, D. (高橋璃子訳) 『GDP——〈小さくて大きな数字の歴史〉』みすず書房
- 小林雅之 (2004) 「大学ランキングの功罪——リングとミカンはどちらがいいか」 『アルカディア学報』2170 号. <https://www.shidaikyo.or.jp/riihe/research/184.html>
- ゴーマン, J. (佐々木一芳訳) (1998) 『ゴーマンレポート——アメリカの大学および世界の主要大学格付け』アイ・エル・エス出版.
- コリンズ, J. (山岡洋一訳) (2001) 『ビジョナリー・カンパニー 2——飛躍の法則』日経 BP.
- (2017) 『ビジョナリー・カンパニー 3——衰退の五段階』日経 BP.
- コリンズ, J.・ハンセン, M. (2012) (牧野洋訳) 『ビジョナリー・カンパニー 4——自分の意志で偉大になる』日経 BP.
- コリンズ, J.・ポラス, J. (1995) (山岡洋一訳) 『ビジョナリー・カンパニー——時代を超える生存の原則』日経 BP.
- コリンズ, J.・ラジャー, B. (土方奈美訳) (2021) 『ビジョナリー・カンパニー ZERO——ゼロから事業を生み出し、偉大で永続的な企業になる』日経 BP.

- 小島寛之 (2006) 『完全独習 統計学』ダイヤモンド社.
- ザイゼン, H. (佐藤郁哉訳) (2005) 『数字で語る——社会統計学入門』新曜社.
- 佐藤郁哉 (2006) 『フィールドワーク 増訂版 書を持って街へ出よう』新曜社.
- (2008) 『質的データ分析法』新曜社.
- (2015) 『社会調査の考え方 [上]』東京大学出版会.
- (2021) 『ビジネス・リサーチ』東洋経済新報社.
- (2022) 『プッシュボタン式統計調査の効用と限界 (1) —— 3 種類の不可解なグラフと改善提案の事例から』『同志社商学』第 74 巻第 1 号, pp.113-144.
- 佐藤俊哉 (2017) 「統計の有意性と P 値に関する ASA 声明」(Wasserstein and Lazar (2016) の “The ASA Statement of Statistical Significance” の部分の邦訳. <http://biometrics.gr.jp/news/all/ASA.pdf>.)
- シュービン, N. (黒川耕大訳) (2021) 『進化の技法——転用と盗用と争いの 40 億年』みすず書房.
- 須藤靖 (2016) 「〇〇のバカヤロ——注目の多い雑文 その 34」『UP』2016 年 6 月号.
- 橘敏明 (1986) 『医学・教育学・心理学にみられる統計的検定の誤用と弊害』医療図書出版社.
- 寺島和夫編著 (2022) 『改訂版 ビジネスデータの分析リテラシーと活用』同分館出版.
- トゥーリッシュ, D. (佐藤郁哉訳) (2022) 『経営学の危機——詐術・欺瞞・無意味な研究』白桃書房.
- 野口武彦 (2005) 「書評委員お薦め『今年の 3 点』」『朝日新聞』2005 年 12 月 2 日付.
- 本田雅一 (2021) 「なぜ「こたつ記事」は増えたのか 10 年前に作った言葉がにわかに注目を集めた理由」<https://www.itmedia.co.jp/news/articles/2101/07/news103.html>
- 三浦展 (2005) 『下流社会——新たな階層集団の出現』光文社新書.
- 三浦展・上野千鶴子 (2007) 『消費社会から格差社会へ——中流団塊と下流ジュニアの未来』河出書房新社.
- 好井裕明 (2006) 『「あたりまえ」を疑う社会学——質的調査のセンス』光文社.
- 米盛裕二 (2007) 『アブダクション——仮説と発見の論理』勁草書房.
- ラインハート, A. (西原史暁訳) (2017) 『ダメな統計学——悲惨なほど完全なる手引き書』勁草書房.
- リクル, P. (久米博訳) 2004 『記憶・歴史・忘却 (上)』新曜社.
- レイ, O. (池畑奈央子・原俊彦訳) (2020) 『統計の歴史』原書房.
- ローゼンツワイク, P. (桃井緑美子訳) (2008) 『なぜビジネス書は間違うのか』日経 BP.
- Allport, G. (1942) *The Use of Personal Documents in Psychological Science*. Social Science Research Council.
- Alonso, W. and P. Starr (1983) “Introduction,” In Alonso, W. and P. Starr eds. *The Politics of Numbers*. Russel Sage Foundation.
- American Council of Learned Societies (1979) *Scholarly Communication: The Report of the National Enquiry*. Johns Hopkins University Press.
- Arnold, D. (1970) “Dimensional Sampling,” *American Psychologist*, 5: 147-150.
- Blumer, H. (1979) *Critiques of Research in the Social Sciences*. Transaction Books.
- Bridgman, P. (1927) *The Logic of Modern Physics*. Macmillan.
- Bulmer, M., Bales, K., and Sklar, K. K. (eds.) (1991) *The Social Survey in Historical Perspective 1880-1940*. Cambridge University Press.
- Bryman, A. and Bell, E. (2015) *Business Research Methods (3rd ed.)* Oxford University Press.
- Cairney, P. (2016) *The Politics of Evidence-Based Policy Making*. Palgrave Macmillan.
- Denzin, N. (1978) *The Research Act: A Theoretical Introduction to Sociological Methods*. McGraw-Hill.
- Desrosières, A. (tr. by C. Naish) (1998) *The Politics of Large Numbers: A History of Statistical Reasoning*. Harvard University Press.
- Duncan, O. (1984) *Notes on Social Measurement: Historical and Critical*. Russel Sage Foundation.
- Edwards, J., and Berry, J. (2010) “The Presence of Something or the Absence of Nothing: Increasing Theoretical Precision in Management Research,” *Organizational Research Methods*, 13, 668-689.

- Ellul, J. (1976) "Problems of Sociological Method." *Social Research*, 43(1) : 6-24.
- Flyvbjerg, B. (2011) "Five Misunderstandings about Case-Study Research" In Seal, C. et al eds. *Qualitative Research Practice*. SAGE.
- Geertz, C. (1983) "Blurred Genres: The Refiguration of Social Thought." In *Local Knowledge: Further Essays in Interpretative Anthropology*. Basic Books.
- Giddens, A. (1979) "reductionism." In G. D. Mitchell ed. *A New Dictionary of the Social Sciences*. Aldine, p.156.
- Knaflig, C. N. (2015) *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Wiley.
- Lundberg, G. (1926) "Case Work and the Statistical Method." *Social Forces*, 5 (1) : 61-65.
- (1939) *Foundations of Sociology*. Macmillan.
- Hazelkorn, E. (2015) *Rankings and the Reshaping of Higher Education: The Battle for World-Class Excellence 2nd ed.* Palgrave Macmillan.
- Medawar, P. (1964) "Is the Scientific Paper a Fraud?" In D. Edge (ed.) *Experiment: A Series of Scientific Case Histories*. British Broadcasting Corporation, Pp.7-12.
- (1967) *The Art of the Soluble*. Methuen & Co Ltd.
- McKinney, J. C. (1966) *Constructive Typology and Social Theory*. Appleton-Century-Crofts.
- Miles, M. B. and Huberman, A. M. (1994) *Qualitative Data Analysis: An Expanded Sourcebook*. SAGE.
- Porter, T. (1986) *The Rise of Statistical Thinking: 1820-1900*. Princeton University Press.
- (1995) *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton University Press.
- Ragin, C. (1994) *Constructing Social Research*. Pine Forge.
- Roth, J. A. (1966) "Hired Hand Research." *The American Sociologist*, 1(4) : 190-196.
- Selingo, J. (1997) "A Self-Published College Guide Goes Big-Time, and Educators Cry Foul." *The Chronicle of Higher Education*, Nov.7.
- <https://web.archive.org/web/20071215111105/http://chronicle.com/free/v44/i11/11a00101.htm>
- Shi, J. C., R. K. Toukoushain, and U. Teichler eds. (2011) *University Rankings: Theoretical Basis, Methodology and Impacts on Global Higher Education*. Springer.
- Singleton, R and Staits, B. 2010. *Approaches to Social Research*. (5th ed). Oxford University Press.
- Smithson, M. (2000) *Statistics with Confidence*. SAGE
- Sorokin, P. (1940) "Foundations of Sociology [book review]" *American Journal of Sociology*, 45(5) : 795-798.
- Starr, P. (1983) "The Sociology of Official Statistics," In Alonso, W. and P. Starr eds. *The Politics of Numbers*. Russel Sage Foundation.
- Sudman, S. (1976) *Applied Sampling*. Academic Press.
- Times Higher Education (2017) *Times Higher Education World University Rankings 2018*. THE.
- Tourish, D. (2019) *Management Studies in Crisis: Fraud, Deception and Meaningless Research*. Cambridge University Press.
- Tukey, J. W. (1962) "The Future of Data Analysis," *Annals of Mathematical Statistics*, 33(1) : 1-67.
- Wasserstein, R. L., Lazar, N. A. (2016) "Editorial: The ASA's Statement on p-values: Context, Process, and Purpose." *The American Statistician* 70(2) : 129-133.
- White, L. A. (1943) "Sociology, Physics and Mathematics." *American Sociological Review*, 8(4), 373-379.
- Yates, F. (1964) "Sir Ronald Fisher and the Design of Experiments." *Biometrics*, 20(2) : 307-321.
- Yudkevich, M., P. G. Altbach, and L.E. Rumbly eds. (2016) *The Global Academic Rankings Games: Changing Institutional Policy, Practice, and Academic Life*. Routledge.