

はじめに

それは、今後三〇年にわたる、英国農業の抜本的かつ意欲的な改革になるはずだった。二〇〇四年二月。マーガレット・ベケット環境・食糧・農村地域大臣は、下院（庶民院）で演説を行い、改革がどのようなものになるのかを説明した。これまで農家はEU（欧州連合）の共通農業政策のもとで、農産物の生産量に応じた補助金を受け取ってきたが、今後は、もつとわかりやすく、「農業・酪農業らくのうに使える土地」の面積に応じて支払いを受けるようになる。この新たな補助金支給制度は公平で環境にもやさしく、しかも面倒な事務手続きも最小限に抑えられる。九カ月のうちに新制度による補助金支給に着手するという「果敢な」スケジュールで始める英国は、EU諸国として初めてこの制度を導入することになるのだ、と。¹

しかし、そのちようど一年後、ベケット大臣は全国農業者組合の組合員を前に、「『深く失望している』という表現では言い表せないぐらい落胆している」と悄然しやうぜんとした口調で語り、「まったく腹が立ってしかたがない」とまでつけ加えた。²

本来ならその時点で新制度への移行は完全に軌道にのっているはずだったのに、まだ始まってさえいなかった。さらにその一年後、当初の予定では移行がほぼ完了している時期であったのに、支払いが済んでいたのは新制度の予算のうちわずか一五%にすぎなかった。³補助金支給の遅れは農家に大打撃をもたらし、一部では自殺という悲劇まで起きていた。一方、この制度導入による予期せぬ費用は、二〇一二年の時点で六億八〇〇〇万ポンド（約八五七億円）にまで膨れ上がり、さらには、農家への補助金を予定どおり支給できなかったことでEUから科せられた罰金は、約一億六〇〇〇万ポンド（約二〇五億円）に上ったのだ。⁴

こうした事態に陥^{おち}ってしまったのは、新制度への移行がどのくらいの規模になるのか、環境・食糧・農村地域省がまったく把握できていなかったからだ。そのおもな原因は、対象となる土地所有者の多さを、同省が想定できていなかった点にある。新制度では、「農業・酪農業に使える土地」のすべてが補助金の対象となったことから、年間申請件数が八〇〇〇件から一二万件に急増した。果樹園や放牧地を所有してさえいれば、栽培や飼育を実際に行っているかどうかとは無関係に、誰でも補助金を申請できたのだ。申請件数は、どこからともなく、まるで存在しない場所からどんどん申請されるかのように増加していった。というのも、「農業・酪農業に使える土地」というデータがそもそも存在せず、もとより、そうした土地のほとんどはそれまで登記されていなかったからだ。結果的に、二〇〇五年の新規登記件数は、通常の約九〇〇〇件に対して一〇万件にも上った。⁵同省は、何が起ころるか、完全的外れな予想をしていたのである。

このようなことが起こったのは、一度だけの話ではない。政府がこの手の見当違いを繰り返すのは、「バッドデータ」【訳注…統計学的に理想的なデータに紛れ込んで分析を邪魔する粗悪なデータ】を利用してゐるからだ。しかも、ほとんどの場合、自分たちがそうした「ヤバイ」データを使っていることに気づいてさえない。誤った情報を故意に流す「フェイクニュース」の存在は広く知られていて、統計データを意図的に歪曲わいさくしたり捏造ねつぞうしたりできることも、誰もがよくわかっている。だが、先ほどの事例では、政府は偽物の統計データにだまされたわけではなく、信頼性の高い公式データを使った結果、それらが不正確で当てにならないことが判明したのだ。本書では、そのような危険な誤情報について取り上げる。

データとはなんだろうか？ 簡単にいえばデータとは情報のことであり、本書ではおもに数値データを指している。つまり、数える、測るといった手法によって得られる情報だ。通常、そうした確固たる数字の裏には、人間の判断という世界が広がっていて、そこには数多くの思い込みに加えて、疑わしい算出方法が共存している場合さえある。また、一般的には偏りがないと考えられている、データ駆動型システム【データに基づいて意思決定するシステム】を信頼しすぎるのも危険だ。

データにまつわる問題点は万国共通であり、それは地球上のどこにいても、何かを数えたり測ったりするときに直面する問題だが、本書ではおもに英国の事例を扱うことにする。英国の政府、マスメディア、世間のデータの使い方に焦点を合わせることによって、政策を決定する際のデータの扱いに関して、英国政府がいつ、どこで、どう誤ったのかを明らかにするのが本

書の狙いだ。

英国の政策決定者は国民に対して説明責任があるため、みずからが下した判断が公正であり、全国民にとって最善であることを証明しなければならぬ。昔の政府は「私たち政府がそう述べているのだから」という考え方によって、なんの問題もなく政策を正当化できていたかもしれない。だが、民主的に選出され、しかも説明責任を負う現代の政府は、根拠のない妄信に頼って行動するわけにはいかない。彼らは、政策が確かな証拠に裏づけられていることを明らかにするよう求められている。それどころか、義務づけられている場合さえある。それゆえ政府は、自分たちが統治する人々についての正確なデータを必要とする。

これはごく当たり前のことに思えるかもしれない。だが、公共政策に関する判断を逸話や風聞、直感、あるいはまったくの思い込みなどに依拠するのではなく、経験的証拠に基づいて下すべく真剣に取り組まれるようになったのは、ここ一〇〇年ほどのことである。データが公正な統治に欠かせないのは、逸話や思い込みなどとは対照的に、データというのは客観的かつ中立的だとみなされているからだ。

この発想の源は、一七世紀末から一八世紀にかけてヨーロッパで興った啓蒙主義けいもうぎにさかのぼる。それまで、世界というのは「誰も理解できず、また理解すべきでもない、一つの大いなる神秘的な謎」だという認識が、当たり前だとされていた。やがてこの考えは、「体系立った注意深い観察を通じて、我々は物事の仕組みを解明し、しかもその仕組みを支配しようとさえできるようになる」という、フランシス・ベーコンをはじめとする思想家たちの急進的な見解に

よって一掃された。

そうして「世界は、測れるもの、理解できるもの、支配できるものだ」というこの発想から、「国民の幸せのために、国は世界を測り、理解し、支配すべきだ」という考えが生まれたのである。この新たな考え方は、「君主や政府は神権に基づいた不可侵な存在などではなく、社会契約を通じて国民に対する責任を負う」ことを意味している。つまりこの契約によって、人は法律に縛られて納税義務を負うことを受け入れ、一方で政府はそれらの法律を施行し、集めた税金を賢く使うという使命を果たさなければならなくなる、というわけだった。

国家統治のための大量のデータ収集が本格的に始まったのは、ヴィクトリア時代【一八三七〜一九〇一年】にさかのぼる。一九世紀初頭は大いなる変化の時代だった。その変化は人々の生活についてのみならず、権力の座にある人々が「我々には貧困、病気、犯罪という悪に対処する責任がある」と考えるようになった点にも表れていた。

それ以来、政府は、現状に関して統計をとる、つまり、数えたり測ったりするという方法を使って物事を定期的に把握しようとする作業を、少しの疑問も抱かずにひたすら続けてきた。一九八〇年代には、統計をとることはもはやあまりに当たり前の作業になり、データをわざわざ公表するまでもないのではないかと言う人も出るほどだった。有力な実業家でサッチャー政権の顧問も務めたデレク・レイナーは、統計データを「サンゴ礁しょうこう」にたとえた。「どちらも、私たちの目を楽しませるといふ理由だけで守られている」というのが、レイナーの言わんとするところだった。

だが、サンゴ礁は単に鑑賞するためのものではなく、海洋の生態系に不可欠なものだ。しかも、嵐の際には防御の役目も果たす。一九九七年に政権の座に就いた英国の労働党は、統計データというのには「民主主義の礎」であり、「優れた公共経営と説明責任のために欠かせない」と表明した。政策決定において、データは単なる飾りではなく、進むべき道を実際に示してくれるパートナーとなった。今日では、政府の政策策定の過程でデータが必要とされるのは、織り込み済みのことだといえる。

もし自分が英国の首相に選ばれたとしたら、あなたなら何をやるだろうか。少し想像してみても、もしかしたらあなたは、手に入れた新たな権力を使って、それまでずっと悩まされてきた問題を解決するかもしれない。あるいは、自分が正しい道だと信じる方向に向かって、国の歩みを進めようとするかもしれない。あるいは、ダウニング街にある首相官邸の部屋の壁紙を新しいものに変えるのを優先するかもしれない。だが、マニフェストで掲げた公約の実現に向けて取り組んだあと、英国という国を治める日常業務をどのように行えばよいのだろうか？ あなたは、実際に何が起きているのかを知りたいと思うはずだ。あるいは、問題を突き止めて解決できるような、英国内に現れる重要な兆候に常に注意を払わなければならないと思うに違いない。そうしたことを実行するには証拠となる情報が必要だ。つまりあなたは、なんらかのデータを入手しなければならない。

政府が政策を打ち出す場合、それが妥当である根拠を示すために、「インパクト評価」というかたちでデータを使うことが多い。インパクト評価とは、予想されるプラス面とマイナス面

を、客観的かつ中立的な観点から数量的に比較検討するといふものだ。この評価を行うには、たとえ現実的にやや無理があつても、対象となる物事を数値化して計算できるようにする必要がある。犯罪の削減や、公共図書館の閉鎖、大気汚染の改善といった問題を、果たして比較評価のための数値で表せるだろうか。人命の救済を数値で表せるだろうか。たとえ困難であつたとしても、公務に携わる人々はなんらかの方法を編み出さなければならず、そのためにデータに頼らざるをえない。

問題は、そのようなときに役立つ「グッドデータ」【統計学的に理想的な良質のデータ】が、常に手に入るわけではないという点だ。たとえば、二〇〇五年のあの農業政策改革において土地の新規登記件数が急増したにもかかわらず、イングランド【英国を構成する四つの国の一つ。ほかはウェールズ、スコットランド、北アイルランド】の土地のおよそ一五%は現在でも未登記のままだ。そのため、土地所有者に関する一元的に管理された記録は存在しない。ちなみに、たとえ正式な手続きを経た地方分権が実現していようと、特別な慣習によつて管轄範囲に大きな偏りがある地方警察が存在していようと、英国の制度の多くは原則的に中央政府の管理下にある。

一般に、物事は重要とみなされるようになって初めて数えられるようになる。だが、それは公共政策の観点からすれば理想的ではない。英国で長年行われてきた、最も総合的なデータ取得方法である国勢調査の結果を見ると、各年代に何が重視されていたかがよくわかる【英国の国勢調査は一〇年ごとに実施される】。たとえば、一九一一年の「出生についての国勢調査」

では、既婚女性は子どもの数について問われた。また、一九五一年の国勢調査では、重要課題であった戦後の近代化への取り組みに役立つよう、住宅事情についての質問が初めて盛り込まれた。その一方で、一九九一年の国勢調査まで、民族意識に関するデータが集められたことはほぼなかった。一八五一年の国勢調査では「教会通い」に関する質問があったが、それ以降二〇〇一年まで、国勢調査で宗教に関して尋ねられることはなかった。奇妙なことに、そのあいだ宗教に関する質問を盛り込むことに反対しつづけていたのは、データによってみずからの存在に疑問が投げかけられるのを懸念したイングランド国教会だった。だが実際、その不安は的中していたといえるのかもしれない。というのも、二〇〇一年以降、四番目に信者が多い宗教は「ジェダイ」だったからだ。【宗教に関する質問への抗議として、映画『スター・ウォーズ』に登場する組織「ジェダイ」と答えようという運動に賛同した人々による回答】。ときとして、「データを集めさせない」という手段が意図的に取られることもあるのだ。

一九八〇年代を通じて、失業は最も重要な政治的課題であったにもかかわらず、当時の失業者数の正確なデータはない。同様に、一九九〇年代から二〇〇〇年代にかけて、政府は犯罪問題に取り組むと約束し、犯罪の減少目標をみずから設定したにもかかわらず、当時の犯罪件数に関する一貫性のあるデータは残されていない。このように、物事をきちんと数えることも、継続して数えることもしていない場合、事態が好転しているのか、それとも悪化しているのかを把握できない。

英国への移入民と、英国からの移出民の数は、長年にわたって把握されてこなかった【国内

への移住者と国外への移住者の区別を特に明確にしたい箇所では、「移民」ではなく「移人民」「移出民」という訳語を充てている。二〇一〇年代初めの公式統計データを元にとすると、英国で学ぶ留学生の半数はビザが切れて不法滞在しているように思われたため、当然ながらその取り締まりが政策の優先事項となった。ところが、「学生ビザは不法滞在者のためのある種の『トロイの木馬』【ギリシャ神話由来で、巧妙な隠れ蓑を意味する】として利用されている」と繰り返し主張してきた一人であるテリーザ・メイ首相にとつては、恥ずべき事態となった。二〇一七年、精度の高い新たな統計データによって、実際の不法滞在者はきわめて少なかったことが判明したのだ。

英国では、不確かな数字を過信し、それに基づいた政策が何十億ポンドもかけて推進されることもあった。あるいは、十分なデータがまだ揃っていない段階で、政策が決定されることもあった。おまけに、特定の産業における就業者数や賃金も正確には把握されていない。国内の各地域の居住者数さえも掴めていない。さらには、居住者のうちの何人が英国人なのかもわからない。

こうした項目について、政府は正確なデータをとるべきなのだろうか。そもそもそれらは数えられるのだろうか。物事のなかには、確立された概念がないために、数えたり測ったりするのが本質的に難しいものがある。したがって、ここ数年のあいだに貧困が改善したか悪化したかを判断するには慎重を要する。ましてや、数十年という期間で比較を行ったり、他国と比べたりするのはもつと大変だ。ただ、よいか悪いかは別として、男女不平等をおもに男女別の平

均収入に基づいて測定するようになった結果、経済と結びついた解決策が示されるようになった。

日刊紙の『デイリー・メール』は、「エチオピアのスパイス・ガールズ」とも呼ばれ、女性の権利拡大推進を目的とする援助金を英国から受けていたガールパワーポップバンドの「イェグナ」を批判する運動を何年も続けていた。結局、同バンドへの援助が二〇一七年に打ち切られたとき、政府報道官は、英国の援助を「もっと有効な方法」で活かすためだと説明した。¹⁰だが、もっと有効かどうかを本当に正しく判断できるのだろうか。「女性の権利拡大」というような、多面的で文化と結びついているものを測って比較するのはほぼ不可能であり、そもそもポピュラーソングを聞くことと女性の権利拡大の直接的な因果関係を示すのも無理なのではないだろうか*。

たいていの場合、明確に定義して測定するのが最も難しいのは、社会的に重要な物事だ。「障害をもっている」や「貧しい」の定義は確立されておらず、「心の病」「孤独」「差別」といった、広く認められた社会問題であっても、満場一致で賛同を得られるような明確な定義はいまなお定まっていない。別の例でいえば、「反社会的行動」については確立された定義もなければ、そういった行動を長期にわたって観測するための定常的な手段もないため、それが実際にどのような変化してきたのかを捉えるのは難しい。それゆえ、世間はマスメディアが与える印象に流されやすくなる。また、いわゆる「刃物による犯罪」の統計データには、「刃がついているか、先が尖っている物」が使用された犯罪がすべて含まれていて、そのなかには「ジャ

ガイモの皮むき器」も入っている。¹¹ まあ、もしあなたがジャガイモだったとしたら、確かにそれは恐ろしい凶器かもしれないが……。

とはいえ、英国においてデータに関する問題が浮き彫りになった最大の事態は、新型コロナウイルス感染症。パンデミック（世界的大流行）の期間だろう。各種のシステムが連動していないため、パンデミックが発生したとき、国民保健サービス（NHS）が管理している記録に個人用防護具（PPE）の在庫数についての情報はなかった。地方自治体レベルでは、社会福祉サービスに従事している人の数に関する記録は存在しておらず、各地域で保健サービスを受けている人の数も正確には掴めていなかった。しかも、どこに誰が住んでいるのかについて一元的に管理された記録がないため、ウイルスに曝露した恐れがある人を探し出して追跡調査するのは困難をきわめた。パンデミックに備えるために最善の対策モデルを構築すべき専門家が、大勢いたにもかかわらず、実際にパンデミックでわかったのは、私たちが何をすべきかについて、統計モデルは明確な答えを出せそうにもないということだった。

加えて、今回のパンデミックがあらためて明らかにしたのは、物事を長期にわたって比べたり、他国と比較したりするのがいかに難しいかということだ。さらには、それでも人々が比べ

* 女性の権利拡大への取り組みが成功しているかどうかを判断するための指標として国際連合（国連）が提案しているものには、「平均収入」「世帯内での資産の分配」「時間の使い方」といった従来の指標に加えて、「管理職に就いている女性の数」「携帯電話を所有している女性の割合（男性との比較）」などがある。

るのに必死になることもわかった。行われている検査の基準はほぼすべての国で異なり、しかも、死亡に関する記録方法もそれぞれ異なっている。それにもかかわらず、自分たちは近隣国と比べてどうなのかという点を常に強く意識させられた。英国内においてさえ、二〇二〇年に「新型コロナウイルス感染症による」死亡数の算出方法が変更されている。さらには、英国内の国家機関に所属する統計職員たちは、例年の死亡数に基づいて、予測を超える「超過死亡数」を分析した結果、「公式の死亡数」が実際の死亡数より少なくとも一万人も低い数字である可能性が高いと結論づけている。¹² また、パンデミック発生時に死亡数を正確に算出する手立てがまったくない国もあった。一般には、中国における死亡数として記録された数字は、実際の半数以下だと考えられている。¹³ 死亡届を出す義務がない国は、アフリカだけでも四六カ国もある。それはつまり、「新型コロナウイルス感染症による二〇二〇年の死亡数はわずか三名」だというブルンジの主張を、額面どおりに受け取るわけにはいかないことを意味する。¹⁴ 『エコノミスト』誌による二〇二二年初めのモデル分析では、全世界の実際の死亡数は各国の公式の数字を集計した合計人数のおよそ四倍と推測された。¹⁵

こうしたことは、何か特別な場合や、専門的な分野でのみ起こる問題ではない。

「バッドデータ」を問題視しなければならないのは、政府が重大な決断をする際には必ずデータを利用するからだ。そして、政策の正当性を確かな証拠に基づいて示すよう求められているために、政府は現実的になや無理があるかたちでも、費用対効果を数値化しようとする。試みる。「バッドデータ」を問題視しなければならない理由はほかにもある。世間は自分たちが見た

聞いたりする情報の出所が国の統計職員といった信頼すべき人々であれば、その情報を信用してしまいがちだ。私たちは、ありとあらゆる物事について、それに関するデータが必ず存在し、手に入れられると思いつく傾向がある。また、善意によるものか悪意によるものかはともかく、「これがあなたの求めているものです」といつて、なんらかの情報を与えてくれる人がいるとき、確かなデータがなければフェイクニュースに反論できる手立てはない。

政治家と統計職員は、物事を数えたり測ったりする難しさと、あらゆることに答えを出すように常に求められている状況のあいだで板挟みになるという困難な事態に直面している。統計職員は、最低限のリソースしかなければできないことが限られ、自信過剰になっても、あるいは逆に念には念を入れすぎても窮地に追い込まれる。統計学上の問題点を解決しようと、研究者たちや経済の専門家たちがどっと押し寄せてくることが多いが、彼らは必ずしも信頼できるわけでもなければ公平なわけでもない。

これは、誰が権力の座に就いたとしてもなくならない問題だ。どんな政府も、政党も、出所や健全性を疑わずにデータを使いがちであり、それを止めるのは難しい。もちろん、「グッドデータ」を集めて賢く活用するために政府は全力を尽くしてはいるが、なかには非常に根深い問題もあり、それらを変えるには多大な努力に加えて投資も必要だ。たとえば政府が、統計データの改善に誠心誠意取り組みたいと思っけていても、いざ実行しようとするとき山のように大きな問題が立ちだかっていることに気づき、政治で解決しようという意気込みははしぼんでしまう。こうした状況とは対照的に、昨シーズンのサッカー・プレミアリーグで、ハリー・ケイン選

手がゴールポストの枠内に左足で何回シュートを飛ばしたかというデータはきちんと存在する。当然ながら、それは政府によるデータではなく、サッカーの試合を分析する業者が集めたデータだ。そうした業者は、イングランドのトップリーグで行われる全試合の内容を手作業で記録するために、大勢の人を雇っている。試合の進行中、記録要員たちはコンピューターに向かい、リアルタイムで表示される試合の様子とピッチの図が重なって表示される画面上をクリックしながら、すべてのパスやタックル、ボールタッチを記録する。そうやって試合の流れが事細かに記録された何百万ものデータポイントは、ブックメーカーやサッカー評論家をはじめとして、お金を払ってでもデータを手に入れた人たちにとっての膨大な分析用データセットとなる。

このように、成し遂げようとする気持ちに加えて多額の資金があれば、入手できるデータの質は驚くほど高くなる。

だからといって、政策のためのデータも、同じぐらいの時間や労力や資金を費やして入手すべきだというわけではない。ただ、そこまで緻密なレベルのデータも、技術的には入手可能だと言いたいのだ。サッカーの試合についてはとても詳細なデータがあるというのに、「有権者の数」や「新型コロナウイルス感染症による死亡数」「犯罪件数の増減」はわからないという事実を前にすると、なんともやりきれない気持ちになる。

ここで、現代社会においてデータがもたらしてくれる可能性に目を向けてみよう。地方自治体や学校、警察への予算の割り当てを決めるための統計モデルやアルゴリズムは、すでに導入されている。そうしたモデルやアルゴリズムは、移民の数の予測、人々の行動の「誘導」、不

正の摘発、経済政策の具体化、犯罪予測でも使われている。また、使用者の位置情報を追跡できるスマートフォンの普及と、誰もがなんらかのかたちでオンラインやデジタルの世界とつながっていることによって、データ利用のまったく新しい可能性の世界が目の前に開かれている。さらには、英国国家統計局（ONS）によると、英国における仕事の七二％は、将来ロボットに取って代わられる可能性が高いという。¹⁶ この予測自体は議論の余地があるかもしれないが、自動化と機械学習の分野における技術革新に、各国政府が大いなる意欲を燃やしているのは明らかだ。

とはいえ、統計モデルやアルゴリズムを用いるそうしたシステムの性能は、設計構想と投入されるデータの質にかかっているため、「データに導かれて正しい答えが手に入る」という発想は、単なる思い込みになってしまいうことも多い。ひどい設計をもとに製造された機械に「バッドデータ」が入力された結果、本来なら対象外の人々に巨額の子算が誤って割り当てられ、それ以外の人々は完全に忘れられるという事態が起きてしまった例もある。「私たち政府がそう述べているのだから」という考え方によって物事を正当化できた日々は、すでに過去のものなのかもしれない。だが、すぐ目の前に迫っている時代、つまり「データがそう示しているのだから」という言葉で正当化される時代も、過去の時代と比べて望ましいものになるかどうかはわからない。

来たる時代を望ましいものにするためには、「グッドデータ」が必要だ。現在使われているデータには問題や重大な欠陥が多数あり、そうした問題を解決しないまま、政策の策定過程で

データがますます大きな役割を占めるような将来を迎えてはいけない。「バッドデータ」のなかでも特にひどいのは、犯罪、移民、収入、社会福祉、失業、貧困、平等といった、人々にとって最も重要な問題に関するデータなのだから。

効果的だと思われる政策について政府が何かを主張するときや、国民にとって最善の策が何かを政府が判断するとき、私たち国民は政府を信頼するよう求められる。もちろん、国民は政府を信頼したいと思っている。そのために彼らを選んだのだから。だが、データに欠陥があるかもしれない、統計にも間違いがあるかもしれないという現実を受け入れてしまったら、私たちはいったい何を信用すればいいのだろうか。裏切られてがっかりしなくてもすむよう、何一つ信じないようにしようと思ったとしても、不思議ではない。とはいっても、私たちは、「あらゆる統計データをはねつける」、または「すべて額面どおりに受け入れる」という、両極端な策のどちらかを選ぶなければならないという事態に追い込まれてはならない。

私たち国民にできるのは、落としどころを見つけ出すことだ。

毎日のようにデータとかかわり、政治家ともかかわっている一員として私が目にしていても、読者のみなさんにも見てほしい。私は、英国下院図書館で統計職員として働いている。

この図書館では、ウェストミンスターにある英国議会のすべての国会議員【英国で国会議員とは下院議員のこと】に対して、議会内での調査サービスを提供している。進行中の議論や審議中の法案に必要な背景説明を用意したり、国会議員から直接質問を受けたりするのがおもな仕事だ。私たちは、すべての政党にサービスを提供しているため、政治的に中立な立場を守

っている。私たちの使命は、揺るぎない事実を歪めることなく活用することにある。

統計職員は日々、「私の選挙区で学校給食が無償で支給されている子どもは何人か？」といったものから、「処分されたアナグマの数は？」「英国の防衛費の額は世界で何番目なのか？」といったものまで、多岐にわたる質問に答えている。政治家は時間の流れの速い世界で活動しているのだから、なんらかのデータが必要になると、ほとんどの場合、「いますぐ」にデータをほしがる。しかも、できるだけわかりやすいかたちのものを求める。そうした要望に対して、「それについてのデータはありません」「データはありますが、注意すべき補足事項もたくさんあります」という答えを、丁寧ながらもはっきりと伝えるにはどうすればいいのか。それを考え出すのも、私たち統計職員のおもな仕事の一部といえる。

本書では、データがどのようにして生まれるのか、データが作成される過程での人間の選択がいかなる根本的影響を及ぼすのか、また、データはなぜ公共政策に不具合をもたらす恐れがあるのかについて解説したい。信頼できて当然だと思っていたデータの裏側を覗けば、みなさんも驚くはずだ。それどころか、ショックを受けるかもしれない。

ここには数学的な話は出てこない。グラフや表もなければ、当然ながら数式も使われていない。というのも、統計データを入手する際の現実的な側面を理解したり、それらのデータが信頼できるかどうかを判断したりするのに、数学の知識は必要ないからだ。大事なものは、データがどんな仮定のもとで得られたものなのかを把握し、人的エラーに気をつけることだ。物事にはどうしても数値化するのが難しいものがあり、その事実を否定するのではなく受け入れる必

要がある。さらにいえば、現在のやり方よりうまく数えたり、測れたりできるようになるものもたくさんある。

データの利用頻度がさらに増え、今後の政策策定においてデータの果たす役割がより重要になると確実にわかってきたいまこそ、データに関するこれまでの問題点を洗い出し、その解決策を考えるべきだと私は考えている。そうした点を踏まえれば、「バッドデータ」や誤った判断に振り回されずにすむようになるはずだ。