

問題

慶應義塾大学総合政策学部・環境情報学部では、2014 年度からの新カリキュラムで「データサイエンス」を、「言語コミュニケーション」「情報技術」「ウェルネス」と並び必須の履修科目と位置づけています。この背景として、現代社会では、様々なデータを収集・分析し意思決定することで、証拠に基づく問題発見・解決の能力を身につけることが、今まで以上に重要視されていることが挙げられます。

両学部のデータサイエンス教育では、学生が自分自身でデータを上手に集めて適切に分析し、様々な意思決定ができるようになることを目指しています。ここでいうデータには、数値だけではなく、文字や映像・音なども含みます。

現実には存在する問題を発見し解決するためには、様々な場面で意思決定が必要となります。しばしば、将来が予見できないとか、想定外の事象が起きるなどといった不確実な状況の下で意思決定を行うことがあります。「不確実さ」を数量的に表現するために確率が用いられます。そのため確率の概念を理解することは社会科学の分野でも重要です。そして意思決定を支援するために、収集されたデータを分析した結果が用いられます。データの分析には、統計学を理解する必要があります。そのため確率と統計は、データサイエンスの根幹となる学問分野として位置づけられています。

この問題は、総合政策学部で学び、研究を行う上でデータサイエンスを修得することの重要性について、みなさんに考えてもらうことを目的としています。

資料1は、17 世紀後半から 20 世紀初頭における、統計学の歴史とその時代背景を説明したものです。

資料2は、20 世紀後半以降の統計学をめぐる状況を紹介したものです。

資料3と4は、同じ統計学者によるものです。資料3では社会科学における「数量的把握」の意味について、資料4では統計学の立場から経済学などの関連した問題について、それぞれ論じられています。

資料5と6は、太平洋戦争開戦時の意思決定とデータとの関係を論じています。資料5は太平洋戦争開戦時に国家の重要政策や物資動員の企画立案を担った当時の企画院の鈴木貞一総裁に対するインタビューです。資料6は、太平洋戦争における旧日本軍の組織としての失敗に関する学際的共同研究を行った成果をまとめたものです。

資料7では、裁判における確率統計の誤用例が紹介されています。

資料1～7を読んだ上で、以下の3つの問に答えて下さい。

問1

資料1～7は、意思決定を行う際に、データを収集し分析することの利点およびその難しさと限界について述べています。資料1～7のなかから重要だと思う資料を2つ挙げて、データを用いることの利点およびその難しさを、それぞれ200字以内で要約し述べて下さい。

問2

以下のボックスに示された社会的課題は、データを収集・分析し、意思決定することにより解決可能と考えられます。

- (1) 生活習慣病の予防や健康管理による医療費の適正化
- (2) 保育所の待機児童解消や男性の育児参加促進による子育て支援
- (3) 選挙区の区割り改定による一票の格差是正
- (4) 外国人観光客の誘致による雇用促進

このうち、あなたが最も関心のある課題を一つ選び、その番号を解答欄①に記入してください。そして、その問題の現状と課題を示すことができる定量的な指標を提案し、その指標の定義を解答欄②に記述してください。ここで、定量的な指標は以下の例を参考にし、例に示された以外の指標を提案してください。

さらに、その指標をあらわす上で必要な資料やデータの収集方法を解答欄③に記述してください。(300字)

定量的な指標の例

例1：人口密度(人/ km^2) = $\frac{P}{A}$

P = ある地域の人口(人)、 A = 当該地域の居住可能な面積(km^2)

例2：高齢化率(%) = $\frac{P_{65+}}{P_{tot}} \times 100$

P_{tot} = 総人口(人)、 P_{65+} = 65歳以上人口(人)

問3

問2で提案した指標を用いて問題解決のための意思決定を行う上で、どのような限界があると思うか、考えを述べてください。(300字)

資料 1

近代科学としての統計学の源流は 17 世紀頃に見られると言ってよい。それは主に、①国家の人口や産業を定量的に把握し表形式にまとめる方法、②社会現象の定量的な観察から大量のデータを要約し一般法則を見出す方法、③確率論を応用し確率的に事象をとらえる方法、という 3 つの流れのなかから発展してきた。この時代は、三十年戦争（1618-1648 年）に引き続く時代である。三十年戦争は、キリスト教のプロテスタント派とカトリック派の宗教戦争と捉えられることもあるが、それは一側面に過ぎず、国家間の覇権争いや旧貴族階級と新興階級との階級闘争として展開された。そのため国家主義的な風潮が学問をも支配し、そのような背景から統計学も生まれてきた。

このうち①の方法は「国勢学（国状学）」と称される。ドイツのヘルマン・コンリングは、地理や歴史、国家の法規から、国家の統治や立場にとって特に重要と思われる要素を抽出し、国家目的と関連付けて整理することを目的とした、「国家に関する知識」（1660 年）という講義を行った。大量観察に基づく定量的な記述というよりも、国家の安寧に関わる顕著な事項を記述するものであった。コンリングの流れを汲むゴットフリート・アッヘンヴァルは、『基礎制度論』（1749 年）の序文の中で、政治家に必要とされる知識の体系という意味で、統計学（statistics）の語源となる Statistik（「国家学」や「国勢学」を意味する）という言葉を用いた。

②の流れは政治算術派とも言われる。英国のジョン・グラントは、当時たびたび伝染病のペストが流行し大量の死亡者が発生していたロンドンで、教会の資料をもとにした死亡統計を分析することにより『死亡表に関する自然的及び政治的諸観察』（1662 年）を著した。その中で多数の人間の集団では死亡年齢や死因に関して一定の法則があることを指摘した。限られた標本データを数量的に観察し法則性を見出すことで、人口推計が可能となることも示した。同じ頃、英国のウィリアム・ペティは『政治算術』（1667 年）を著し、政治・財政の諸問題を実証的・定量的に検討し、社会現象を数量的に把握することにより規則性を見出す方法を確立した。プロイセンのヨハン・ズースミルヒは『神の秩序』（1741 年）の中で、小規模なデータでは規則性を見出すことは難しいが、大規模データの収集により規則性が見出されるということを指摘している。

社会的な現象を大量に観察することによって規則性が見いだされるという事実は、「大数の法則」とよばれ、その発見は統計学の発展に貢献した。しかしその根拠を明らかにするには、③の確率論の発達が必要であった。ヤコブ・ベルヌーイが 1713 年に『推計論』を著したが、この中に大数の法則の証明が含まれている。英国では長老派の牧師で数学者のトーマス・ベイズが条件付き確率に関する「ベイズの定理」の礎となる概念を発見した。これは後にフランスの天文学者で数学者のピエール・シモン・ラプラスによって定理として完成された。ラプラスは『確率の解析的理論』（1812 年）で数学的確率を理論的に体系化した。彼はフランス革命下の政治家としても活躍し、短期間ナポレオン・ボナパルトの閣僚をつとめた。しかし皇帝ナポレオンが退位した 1814 年版の同書では、「広範な統治を熱望する皇帝たちの衰退は、確率の計算法に精通した者により、非常に高い確率で予言

することができる」とナポレオンを痛烈に批判している。

ベルギーの天文学者ランベール・アドルフ・ケトレーは 1835 年に『人間とその諸能力との発達について—もしくは社会物理学試論』を著し、人間の行動も社会現象も物理学的な法則に従うという考えを述べたほか、身長・体重や道徳観が平均的な「平均人」が理想であるとした。ケトレーはまた、官庁統計データを活用して社会現象に法則性を見出す研究を行った。ここには、①国勢学と②政治算術の融合及び③古典確率論の応用が見られる。彼は官庁統計を発達させ、1846 年にはベルギーで科学的な国勢調査を実施するなど、「近代統計学の父」と言われている。

この頃から、世界各国で国勢調査が定期的な実施されるようになった。国勢調査はセンサスとも呼ばれる全数調査であり、原則としてある時点に国内にふだん居住しているすべての人口や世帯を対象に、その状態を把握することを目的としている。アメリカ合衆国では既に 18 世紀末に全国的な国勢調査を実施しており、19 世紀に入るとヨーロッパ諸国で次々に国勢調査が実施された。日本では、1902 年に「国勢調査ニ関スル法律」が公布されたが、日露戦争や第一次世界大戦の影響により、第一回国勢調査が実施されたのは 1920 年のことである。国勢調査の調査対象者や調査項目は各国の実情に応じて異なるが、初期の国勢調査では調査項目が少なかった。

当時の統計学では、統計データの規模が大きければ大きいほど良いという大標本主義が重視されていた。カール・ピアソンは、ヒストグラムや標準偏差、相関係数などの基礎的な概念を統計学に導入し、データから現象を記述する「記述統計学」をつくり上げるとともに、観測データからデータの分布を推定する方法を提案するなど、近代統計学における統計的推測の基礎を築いた。

あらゆる分野で大規模データを収集することは、当時は困難な事も多かった。ギネスビール社の技師だったウィリアム・ゴセットも、小規模標本データしか得られなかった統計家の一人であり、その経験から『平均の確率誤差』（1908 年）を著した。ゴセットに触発された英国の農事試験場の統計学者ロナルド・フィッシャーは、あるデータ全体から抽出された標本データから、標本データの分布を介して元のデータ全体の特徴を推測しようとする「推測統計学」の方法を導出し体系化した。小標本主義の下での「推測統計学」は、様々な方法論の数理的定式化を厳密に行うことから、「数理統計学」とも呼ばれている。

こうした時代背景から、明治期の日本でも「統計学ブーム」が起きた。福澤諭吉は『文明論之概略』で 18 世紀後半から実用化され始めた統計学に「スタチスチク」と原語を音訳し、その教育研究を奨励した。箕作麟祥は 1874 年にモロー・ド・ジョンネの『統計学』を翻訳し、「政経国表」などの字があてられていた「スタチスチック」を「統計学」と訳した。

参考文献：（一部編集・改変）

武藤徹著（2012）『大量現象のはなし【確率・統計篇】』，日本評論社。

資料 2

太平洋戦争以前の日本でも公的な統計は整備されていたものの、統計の基本法規が未整備であるだけでなく、政府統計の質の悪さや重複調査の氾濫、省庁間での政府統計の調整機能の不備などといった問題が指摘されていた。国家の重要政策や物資動員に関して、データに基づく意思決定が必ずしもなされないこともあった。太平洋戦争下で戦時体制が強化された日本では「客観的な認識のために必要な正確なる統計を欠き、或いはそれが国民の眼からかくされていたために国の基本政策を誤らしめ、国民に戦争の惨禍をもたらした」と、統計法（1947 年制定）の立法の趣旨の中で指摘されている。

日本における戦後の統計制度の改変は、太平洋戦争後に米国大統領府から派遣された統計使節団によるライス勧告を受けたことによる。この勧告は、太平洋戦争時の政府が真実のデータを隠し誤った方向に導いたことを批判した上で、日本の再建と民主主義の確立には、国民が事実にもとづき客観的・合理的に判断し、国の方向性を決定できるような統計への認識と理解を持つことが必要であることを指摘している。さらに、日本の統計機構の再建と統計教育の充実も提唱している。

現代日本では、中央省庁や地方自治体が独立して統計情報を集約し、国勢調査など公的統計の大規模な調査が定期的に実施されている。近年では、公的統計の当初の目的外での二次的利用や匿名データの利活用も行われている。国勢調査の結果は、日本では衆議院小選挙区の画定基準、地方交付税交付額の配分、将来の少子高齢化や人口の見通しといった国や自治体の政策だけでなく、民間企業の経営判断や学術研究にも用いられている。

20 世紀後半以降、コンピュータが発達・普及したことで、統計データ分析と情報処理との関係がより密接となる。第一に、「ビッグデータ」と呼ばれる大規模なデータを扱うことが容易となった。大規模なデータや、様々な分野のデータを組み合わせたデータを活用することにより、新しい知識を得られると期待されている。第二に、先進諸国で問題発見解決の統計教育カリキュラムやガイドラインの整備が進められた。日本でも、2011 年から 2013 年にかけて段階的に改訂された学習指導要領において、統計教育の拡充と必修化が実現した。小学校『算数』では「数量関係」の領域が新設され、小学校高学年で平均や分散について統計的に考察する内容が扱われている。中学校『数学』では「資料の活用」が新設され、高等学校『数学』では「データの分析」が数学Ⅰの構成内容の一つとなった。高等学校『情報』の学習指導要領では、情報社会における問題解決の手法として統計手法の活用が挙げられている。

第三に、条件付き確率に関するベイズの定理を用いた「ベイズ統計」への関心が高まった。ベイズ統計ではある事象（結果）が起きる前提となる事象（原因）が生じる確率を事前確率（主観確率）と呼び、ある事象が起きた場合の「条件付き確率」を事後確率という。ベイズの方法とはいわば、

事前の直感に基づく判断と反復可能な実験に基づく確率を組み合わせ、当初の考えを客観的な新しい情報に基づいて部分修正することにより、事実の観察から原因にたどり着くという方法である。この方法は、不確実な状況の下での確率的な意思決定を行う際には有用であると早くから認識されていたが、事前の確率を主観（経験）により決定してしまう恣意性が多い統計学者から敬遠されてきた。コンピュータ・シミュレーションにより事後の確率が事前の確率に影響を受けないようにする方法などが提案されたことで、ベイズ統計は様々な分野で応用されている。

現代では法廷でも統計学のコンサルタントが活躍し、確率統計の取り扱い方がしばしば取り沙汰される。それは、データの収集と分析による統計的証拠が、科学的証拠の一つとして考えられるからである。米国では、1923 年に下した判決において「専門家証言の根拠はその専門分野での一般的承認を得たことが十分立証されなくてはならない」と科学的証拠に対する基準を初めて明確にした。しかし 1975 年には、「専門家証人が意見や推論形成に際して合理的に依拠した特定分野における事実やデータ自体に証拠能力は不要である」とし、従来の科学的証拠に対する基準を曖昧にするような連邦証拠規則が示されている。また英国では 2011 年に、条件付き確率に関するベイズの定理を用いて犯人を特定する確率を予測する方法を刑事裁判に持ち込むべきでないとの判決が下された。

参考文献：（いずれも一部編集・改変）

国友直人・山本拓監修（2008）『21 世紀の統計学 社会・経済の統計科学 I』，東京大学出版会。
渡辺美智子・椿広計編著（2012）『問題解決学としての統計学』，日科技連。
シャロン・バーチュ・マグレイン著・富永星訳（2013）『異端の統計学ベイズ』，草思社。
田中規久男著（2013）「米国における統計的証拠の意義と技法」，法政論集, 250 号, pp.347-371.

資料 3

1 はじめに

社会科学の各分野において、数字や数式を用いて議論を進めることは、近頃ますます盛んになりつつある。ウィリアム・ペティは『政治算術』の中でいっている、「単に比較級や最上級の言葉を使用したり、また思弁的な議論をする代わりに、わたしのいおうとするところを数と重量と尺度によって表現するという方法」は、「感覚に訴えることのできる議論のみを用い、また自然の中に実見し得る基礎をもつような諸原因のみを考察する方法であって、個々人の移り気や好みや感情に左右されることではない」。

そうしてペティが上述の『政治算術』を書いた一七世紀後半とは違って、現在では、このような方法は、「いまのところでは、あまりありふれたものではない」どころか毎日の新聞雑誌から、選挙演説にいたるまで、きわめてありふれたものになっている。

しかしながら、逆に数字や量的な尺度を用いることによって、議論に見せかけ上の正確さを与えて、人を誤らせる危険もまた決して少なくないこともわれわれが日常経験しているところである。

従って「数と量と尺度とを用いる方法」の意味と限界とを吟味してみることも、現在ではきわめて重要な課題であると思われる。〔中略〕

社会科学における数量的方法は一定の限界を持ち、その限界内では一定の有効性をもつなどといってみても、それだけでは無意味である。それよりもまず、社会科学における数量的認識のあり方について、立ち入って考えてみなくてはならない。

2 量の客観性

社会科学において（あるいは自然科学においても）数量を用いることの効果は、すでにペティが指摘しているように、第一にそれによって命題が客観的な形で表現されるという点にあるといわねばならない。ここで客観的というとき、それは二重の意味を持っている。すなわち一つはそこにいわれている趣旨の意味が、客観的にはっきりした意味をもっているということである。「日本には人口が非常に多い」という代わりに、「日本は一億三〇〇万人の人口をもつ」といえば、それがどのような事実に対応しているかという点で、その意味するところはきわめて明白になる。「客観的」というときの第二の意味は、命題の内容が、他の命題との関連が疑問の余地のない形で与えられるということである。「人口が非常に多い」というのは「かなり多い」より大きい人口を意味しているのは明らかであっても、どれだけ多いのかは明らかではない。これに対して「一億三〇〇万の人口」は例えば「四八〇〇万の人口」より「五五〇〇万多く」「ほぼ二倍である」ことは明らかである。

ところで数量による表現が二重の意味で客観的であることに応じて、それは現実に二重の効果を持つ。第一に少なくとも原理的には、命題の真偽を経済的に検証することが可能であるということである。「日本の人口が一億三〇〇万である」かどうかは国勢調査の数字を参照すればチェックできる。第二に、数量的表現をとることによって命題の間の論理的関係が明確にされる。例えば「毎年1%ずつ人口が増加すれば、一〇〇年後に日本の人口がどれだけになるか」は単純な計算ですぐ

導くことができる。従って「一九七〇年の人口が一億三〇〇万である」「人口は毎年 1%ずつ増加する」「二〇〇〇年の人口は一億二〇〇〇万である」という三つの命題が同日成り立つことはないということはすぐわかる。従ってもし最初の命題が事実として正しいならば、あとの二つのうち少なくとも一方は誤りでなければならない。この場合、どちらが誤りであるかを事実と照して直ちに検証する必要はない。ただ全体の体系が再検討を必要とするということを明らかにしただけで、十分意味のあることが多い。

数量的方法の効果は、主として第二の意味の客観性によることが多い。実際、数量というものは、多くの場合それだけでは意味をもたないことが多い。人口が一億三〇〇万であるということにしても、それが五年あるいは一〇年前と比較してどれだけ増加したか、あるいは日本の国土面積と比較してどうか、また他の国と比較すると、例えば中国の何分の一かというようなことを考えて、はじめて意味のあるものとなるので、それだけ孤立した数字として取り出しても意味がない。

量が、いわば関係を明らかにするのに有効であるということは、数量そのものが具体的に与えられなくても、あるいはきわめて大まかな近似値としてしか与えられなくても、いろいろな議論を進めるには十分であることも少なくないということを意味する。例えば人口が一億三〇〇万というのは日本の一九七〇年の人口の概数であるが「一九七〇年の日本の人口」というものもそれだけではあいまいな概念である。一九七〇年といっても年の始めと終わりでは人口には変化があるし、また「日本の」人口とは何か、についても細かくいえば無限に多くの問題がある。国勢調査の場合には、「一九七〇年一〇月一日午前零時に日本国内にいる人の数」というように出来る限り細かく規定しているが、それでもいろいろな問題は残るのみならず、他方ではいろいろな誤差や調査漏れが生ずることが避けられないのであって、人口の正確な値などともいうような一見簡単なものでも本当にそれを求めることはほとんどできないといわねばならない。しかし実は本当に正確な値というものの必要ではない。いろいろな関係を考えるには、例えば約一億三〇〇万という値が知らればそれで十分であって、それが実際には一億二九〇万でも一億三二〇万でもほとんどさしつかえない。〔中略〕

実際、数学的論理の特徴は、対象そのものの「存在」よりも、対象の間の形式的な関係、ないし「構造」を重視するところにある。近代的な数学においては、「数の本質は何か」というようなことは問わない。ただ数の間の形式的な関係を公理として前提して、それから論理的にいかなることが導かれるかを追求するのである。その展開のプロセスは形式的な論理の規則さえ知っていれば誰にでもわかるという意味で「客観的」であり、またその形式はいかなる現実の実体とも結びつけられてもいないし、従って逆にどんな対象についても、同じ構造をもつ限りあてはめることができるという意味で「普遍的」である。

そこで数量を用いることの有効性と妥当性は、対象にどこまでこのような形式的な関係をあてはめることができるかという点にかかっていることがわかる。〔中略〕

3 量の実体

まず第一に、数量というものを具体的に計測するということと、数量それ自体とは区別して考えねばならない。数量というものは一応その観測値とは独立のものとして考えねばならない。〔中略〕

量というものがいわば完全に成立しているためには、単に等しい量のものは同じであるというだけでなく、それがまた同質の部分から成っており、その一部を他の量の等しいものでおきかえることができるということが成り立たねばならない。〔中略〕物質の量的な把握が成立したとき、すべての物質の相違は「本質的には」量のみの差で説明されるべきものであると考えられたのであった。〔中略〕

このような量は、勿論何らかの意味で観測可能なものであるか、観測可能なものと結びつけて理解されうるものでなければならない。観測によつては全くチェックできないような量を考えることは無意味である。しかし直接観測可能でないということは、直ちに量の存在を否定することにはならない。そうでなければ現在、直接人類の手の届かないところにある物質の質量を考えることも許されなくなってしまうであろう。しかしまた逆に数値に表れるかたちで観測可能なものが、つねに量を表わしているとは限らない。いわば量の存在しないところでも観測可能な事実と数値とを客観的な方法によつて結びつけることは可能である。

出典：竹内啓著（1971）『社会科学における数と量』，東京大学出版会。（一部編集・改変）

資料4

統計学の性質

統計学者の大部分は、統計学を固有の対象を持った一つの科学とは考えていない。そこで統計学は、しばしば「方法科学」である、或いは「方法的補助科学」であるなどと言われている。私はこのような言い方に反対でもないが、「方法」という言葉で、対象と無関係に作られる一つの数学的体系、或いは社会科学と自然＝技術科学のそれぞれに対応して考えられる二つの体系というように理解されるならば、それは正しくないと思う。

私は統計学というものは、人間であれ自然であれ、客観的对象に関する数字データの性質を明らかにし、その意味するところを分析する方法の、「ゆるい連合体」であると思っている。すなわちそのような方法は、対象の性質や分析の目的に制約されるが、同時に異なる対象について同じ方法が有効である場合もあるのである。したがってそれはひとつの体系にまとめることはできないが、そこに含まれる諸方法の中には、対象を分析する際の共通な観念というものがあり、対象に応じて全くばらばらに分かれてしまうものでもないのである。私はその共通な観点を「統計的なものの見方」と呼んでいる。いいかえれば統計学とは、「統計的なものの見方」を具体化し、かつその意味を論じた諸方法、諸理論の集まりであるといってもよいであろう。

「統計的なものの見方」は体系的ではない。それはつねに現実を与えられたデータから出発するのであって、抽象化された要素概念を作り出すことではない。〔中略〕

物理学や経済学の場合、その理論は、少なくとも建前としては、対象の客観的性質を、それを観察したり定式化したりする人間とは無関係に記述したものとなっている。ところが統計学の対象となる統計的方法は、データが得られた対象の構造についてだけでなく、そのデータを取り、かつ特定の目的に利用する主体の側の観点を導入しなければ、論ずることができない。どういう統計的方法が適切であるかは、客観的对象のあり方だけでなく、それから得られたデータをどのような目的にどう利用するかにも依存している。統計学がこのような場における「決定のルール」を対象とする、先にのべた統計的決定理論の考え方は、この点に根拠を持っているのである。しかしながら他方、現実の数字データの種類と、その利用のしかた、目的は当然のことながら種々さまざまであるから、それを統一的に体系化することが不可能であることも、ほとんど自明であろう。

統計学の対象となる数字データは、何よりも具体的な対象について具体的な観察の結果として得られた値であり、それはもちろん対象の性質を何らかの形で反映しているけれども、それが直ちに対象の量的性格を表すわけではない。それは観察や分類というような、ある意味で人為的な操作を加えて作られたものであり、そういうものとしての存在理由と、また限界を持っているのである。

〔中略〕

学問の「志向」の方向

統計学や、統計学者は、実は経済学や経済学者とだけ協力するわけではない。私の個人的経験によっても、〔中略〕工学、農学、生物学、医学、心理学、社会学、政治学等、ほとんどすべての学部

に属する人々と、何らかの形で共同研究を行なったり、特定の問題について議論をしたことがあるが、そこでいろいろな分野の間の「ものの見方」の違いを、興味深く感ずることが多い。〔中略〕

研究者がある現象に接したとき、それを見る見方に、「原因指向的」「結果指向的」「過程指向的」の三つのタイプがあるように思われる。「原因指向的」というのは、ある現象に対して、それがどういふところに由来するか、その原因あるいは本質などにもつぱら関心を向け、それを明らかにすることに努力を注ぐという態度である。これに対して「結果指向的」というのは、その現象が何を生み出すか、結局のところどういふことになるのかということにもつぱら注意して、その予測と結果の利用に努力する態度である。また「過程指向的」というのは、現象のなりゆきそのものに興味を持ち、できる限りくわしくそれを知ろうとつとめる態度である。〔中略〕

統計的方法は、元来「データはいかなる事実を示しているか」に関心を持つものであり、「統計的なものの見方」は「過程指向的」経験主義である。〔中略〕

統計学者の感覚

統計的方法の「過程指向的」経験主義的性格は、一方ではその限界を表している。統計的方法によってその名に値する科学的「法則」が発見されたことはないし〔中略〕これからもそういふことはないであろう。しかしこのことはつねに欠点であるとは限らない。早まって一般的法則性を求めようとするのが、不毛な努力に終わることも少なくないからである。〔中略〕たとえばいくつかの経済量の間の関係がデータから知られた場合、経済学者は、そのような関係を説明する「理論モデル」、すなわち企業や家計の行動についての「公理」から出発して、観測された関係式を導き出すような数学的論理を作り出すことを目指す。これに対して統計学者はそのような関係がどの程度安定的であるか、あるいはどの程度の範囲で成立するかについて、まず関心をもつ。経済学者の目から見れば、統計学者は「没理論的」な「悪しき経験主義者」のように見える。〔中略〕

しかし事実問題として、確かに「没理論的」であることは好ましいことではないけれども、経済学の「理論モデル」を現実にあてはめるのに必要な諸前提の検討をおろそかにして、早急な理論化を行うことは危険であることはいうまでもないし、またそういう理論化ができない限り、事実上存在する関係も無視するということが誤りであることも当然である。〔中略〕

経済学者と統計学者の相違は、単に理論の妥当性をどれだけ信じるかという点にあるわけではない。より根本的な「ものの見方」の違いは、経済学者は「理論モデル」に到達することによって、「現象」の背後にある「本質」に到達したと感ずるのに対して、統計学者にはそれは現実の関係が成立する一つの説明を与え、それによってその妥当性を相対的に高めうるものとしか感じられないのである。この二つの態度は、互いに優劣を争うべきものではないであろう。それはそれぞれの学問の性格の問題である。

出典：竹内啓著（1977）『統計学と経済学のあいだ』，東経選書。（一部編集・改変）

資料5

—（注：昭和十六年）十一月五日の御前会議の報告はどういう気持でしたか。

「僕は腹の中では、アメリカと戦争をやって勝てるとは思っていなかったから、とても憂鬱な気持ちで読み上げましたよ……。あの時はねえ、陸軍が戦争をやるといっていたが、実際にアメリカとやるのは海軍なんだ。海軍が決心しないとやれない、陸軍は自分でやるんじゃないから腹がいたまない、それで勝手なことをいっていたんです。海軍は自分がやるんだから、最終的な決断は海軍がすべきだったんだ。ところが海軍は、できないとはっきりいわんのだ」

数字というものは冷酷だと、しばしばいわれる。数字は客観的なものの象徴であり、願望などいっさいの主観的要求を排除した厳然たる事実の究極の姿だと信じられているからである。数字がすべてを物語る、という場合、それはもはや人智を超えた心理として立ち現れている。数字は神の声となった。

しかし、コンピュータが、いかに精巧に作られていても、データをインプットするのは人間である、という警句と同じで、数字の客観性というものも、結局は人間の主観から生じたものなのであった。

鈴木は、自分の報告の効果を知っていたはずである。

— 企画院総裁の提出した（注：戦争遂行に必要な石油などの物資に関する）数字は「（注：戦争をやる）ためのつじつま合わせに使われたと思うが、その数字は「客観的」といえますか。

「客観的だよ。戦（いくさ）にならないように、と考えてデータを出したつもりだ」

— でも石油は南方進出した場合にのみ「残る」とでていたが……。

「戦争を何年やるか、という問題なんだ。仕掛けたあとは緒戦に勝利して、すぐに講和にもっていく。その戦はせいぜい一年か二年。そうすれば石油は多少残る、と踏んでいたんだ」

— しかし、三年分の数字が提出された。

「……とにかく、ぼくは憂鬱だったんだよ。やるかやらんかといえば、もうやることに決まっているようなものだった。やるためにつじつまを合わせるようになっていたんだ。僕の腹の中では戦をやるという気はないんだから」

— 「やる」「やらん」ともめている時に、やる気が無い人が、なぜ「やれる」という数字を出したのか。

「企画院総裁としては数字を出さなければならん」

— 「客観的」でない数字でもか。

「企画院はただデータを出して、物的資源はこのような状態になっている、あとは陸海軍の判断に任す、というわけで、やったほうがいいとか、やらんほうがいいとかはいえない。みんなが判断できるようにデータを出しただけなんだ」

— 質問の答えになっていないと思うが、そのデータに問題はなかったか、と訊いているのです。

「そう、そう、問題なんだよ。海軍は一年たてば石油がなくなるので戦はできなくなるが、いまのうちなら勝てる、とほのめかすんだな。だったらいまやるのも仕方ない、とみんなが思い始めてい

た。そういうムードで企画院に資料を出せ、というわけなんだな」〔中略〕

「やる」という勢いが先行していたとしても、「やれる」という見通しがあったわけではなかった。そこで、みな数字にすがったが、その数字は、つじつま合わせの数字だった。

いわば、全員一致という儀式をとり行うにあたり、その道具が求められていたにすぎない。決断の内容より“全員一致”のほうが大切だったとみるほかなく、これがいま欧米で注目されている日本的意思決定システムの内実であることを忘れてはならない。

出典：猪瀬直樹著（2010）『昭和16年夏の敗戦』，中公文庫．（一部編集・改変）

資料 6

主観的で「帰納的」な戦略策定—空気の支配

戦略策定の方法論をやや単純化していえば、日本軍は帰納的、米軍は演繹的と特徴づけることができるだろう。演繹をある既知の一般的法則によって個別の問題を解くこと、帰納を経験した事実のなかからある一般的な法則性を見つけることと定義するならば、本来の戦略策定には両方法の絶えざる循環が必要であることはいうまでもない。しかしながら、両軍の戦略策定の方法論の相違をあえて特色づけるならば、上記のような対比が可能であろう。さらに厳密にいうならば、日本軍は事実から法則を析出するという本来の意味での帰納法も持たなかったとさえいえるべきかもしれない。

日本軍の戦略策定は一定の原理や論理に基づくというよりは、多分に情緒や空気が支配する傾向がなきにしもあらずであった。これはおそらく科学的思考が、組織の思考のクセとして共有されるまでには至っていなかったことと関係があるだろう。〔中略〕

もともと、科学的な数字や情報、合理的な論理に基づく議論がまったくなされないというわけではない。そうではなくて、そうした議論を進めるなかである種の空気が発生するのである。〔中略〕

日本軍は、初めにグランド・デザインや原理があったというよりは、現実から出発し状況ごとときには場当たりの対応し、それらの結果を積み上げていく思考方法が得意であった。このような思考方法は、とりわけ不確実な状況下において、きわめて有効なはずであった。しかしながら、すでに指摘したような参謀本部作戦部における情報軽視や兵站軽視の傾向を見るにつけても、日本軍の平均的スタッフは科学的方法とは無縁の、独特の主観的なインクリメンタリズム（積み上げ方式）に基づく戦略策定をやってきたと言わざるをえない。

出典：戸部良一・寺本義也・鎌田伸一・杉之尾孝生・村井友秀・野中郁次郎著（1991）『失敗の本質—日本軍の組織論的研究』，中公文庫。（一部編集・改変）

資料 7

ベイズの理論は、B が起きる場合に A が起きる確率（注：条件付き確率）は、A が起きる場合に B が起きる確率とは異なることを示している（注：前者を論じる際に後者を論じたり、後者を論じる際に前者を論じたりする際に生じる過ちを、筆者は「あべこべの誤謬」と呼んでいる）。〔中略〕

法律の世界では、この「あべこべの誤謬」は、ときおり「訴追者の誤謬」と呼ばれている。なぜなら訴追者はしばしばそういった誤った議論をもち出し、それにより陪審員たちは乏しい証拠にもとづいて容疑者に有罪の判決を下してしまうからだ。たとえば、イギリスにおけるサリー・クラーク事件を考えてみよう。クラークの第一子は生まれて 11 週間で死んだ。その死は乳幼児突然死症候群（SIDS）によるとされた。この診断は、乳幼児の死が予期しないもので、しかも検死により死の原因が明らかにならないときになされる。クラークはふたたび子供をもうけたが、今度は 8 週間でその子が死に、ふたたび SIDS によるとされた。

ところがその子が死んだとき、二人の子供を窒息死させたとして告訴され、彼女は逮捕された。裁判で検察側は小児専門医のロイ・メドー卿を呼び、SIDS がめったに起きない病気であることを根拠に、二人の乳幼児が SIDS で死ぬ確率は 7300 万分の 1 であることを立証した。検察側は、彼女に対してほかにこれといった重要な証拠を提示しなかった。はたしてそれだけで有罪とするには十分だっただろうか？陪審員団は十分であると考えた。そして 1999 年 11 月、クラーク夫人は刑務所へ送られた。

メドー卿は、一人の乳幼児が SIDS で死ぬ確率は 8543 分の 1 であると評価していた。彼は各乳幼児に対するその数を二つ掛け合わせることで、7300 万分の 1 という評価をはじき出したのだ。しかしこの計算は二つの死が独立していると仮定している。つまり、第一子が SIDS で死んでいても、第二子のリスクを高めるような環境的或いは遺伝的影響は作用していないと仮定しているのだ。実際、その裁判の数週間後、『ブリティッシュ・メディカル・ジャーナル』誌の論説では、二人の兄弟が SIDS で死ぬ確率は 275 万分の 1 と見積もられた。しかし、それでもまだ非常に小さい。

サリー・クラークが誤って刑務所送りになったことを理解する鍵は、ふたたびあべこべの誤謬を考えることだ。われわれが求めるべき確率は、二人の乳幼児が SIDS で死ぬ確率ではなく、死んだその二人の乳幼児が SIDS で死んだ確率である。クラークが刑務所に送られてから二年後、英・王立統計学会はプレスリリースでこの問題を比較検討し、陪審員の決定が根拠にしたのは、「訴追者の誤謬として知られる深刻な論理の誤りである。陪審員団は、乳幼児の死に対する二つの競合する説明を比較して評価する必要がある。すなわち、SIDS か殺人か、を。SIDS による二度の死であれ、二度の殺人であれ、どちらも極めて蓋然性が低いが、この場合明らかにそのうちの 하나가起きている。重要な問題はその死の相対的な蓋然性であって……（SIDS の説明が）どれほど蓋然性が

ないか、ではない」と断じた。

ある数学者がその後、ひとつの家族が SIDS または殺人によって二人の乳児を失う相対的蓋然性を評価した。彼は利用可能なデータにもとづき、二人の乳児は殺人の犠牲者であるとするより、SIDS の犠牲者であるとするほうが、9 倍蓋然性が高いと結論づけた。

クラーク家は上訴し、その上訴のために、専門家の証言者として独自に複数の統計学者を雇った。クラーク家は上訴審で負けたが、死に対する医学的説明を求めつづけ、その過程で、第二子は死ぬとき細菌感染にかかっていたという事実を、検事側の病理学者が抑え込んでいたことを暴いた。第二子の死はその感染によってもたらされた可能性があった。この新しい事実をもとに、裁判官が有罪判決を破棄し、サリー・クラークはほぼ三年半の刑務所暮らしから開放された。

出典：レナード・ムロディナウ著・田中三彦訳（2009）『たまたま一日常に潜む「偶然」を科学する』，ダイヤモンド社。（一部編集・改変）

2015(平成27)年度 総合政策学部 一般入学試験問題 訂正

教科・科目	ページ	設問	誤	→	正
数学および外国語	10 ・ 11	Ⅱ	問題文 図の各線分は・・・ * 図が問題冊子に掲載なし	→	11ページに以下の図を追加 
数学および外国語	15	Ⅳ	問題文 ・・・ 1 を選択する場合, 解答用紙の V-1 をマークし, 2 を選択する場合, V-2 を・・・	→	問題文 ・・・ 1 を選択する場合, 解答用紙の <u>Ⅳ</u> -1 をマークし, 2 を選択する場合, <u>Ⅳ</u> -2 を・・・
小論文	3	問2	問題文 ・・・意思決定することより・・・		問題文 ・・・意思決定することにより・・・
小論文	9	資料3	上から2行目 ・・・が同日成り立つ・・・		上から2行目 ・・・が <u>同時</u> に成り立つ・・・
小論文	9	資料3	上から21行目 ・・・正確な値というものの必要・・・		上から21行目 ・・・正確な値というものは <u>必要</u> ・・・
小論文	17	資料7	上から9行目 ・・・から開放された。		上から9行目 ・・・から <u>解放</u> された。