

会長講演

歴史の転機と統計学の未来

竹 内 啓

1995年になって、大きな事件、阪神大震災、地下鉄サリンーオウム真理教事件、急激な円高などが相次いで起こっている。戦後50年を迎えて日本の社会も大きな転機に立っているように思われる。

世界的には、1989年のベルリンの壁崩壊の後、ソ連東欧の社会主義政治経済体制が急速に解体し、冷戦は終結した。そのころから、アジア諸国の経済発展が目覚ましく、その勢いは潜在的な超大国である中国とインドも巻き込みつつある。

20世紀の末になって、世界の政治経済状況は大きく変わりつつある。その中で人口の問題、環境問題などが、全人類にとっての重大な問題として明らかとなりつつある。

このような日本全体、人類全体にかかわる「大問題」は、我が日本統計学会のような小さやかな団体にとっては、あまり関係ないと思われるかもしれない。しかしそれは、やはり私達の学問的関心にも大きな影響を与えずにはおかないはずである。

20世紀末から21世紀に向かって、日本も世界も新しい時代に入りつつある現在、統計学は現代の課題に対してどのような有意義な関係 Relevance を持ち得るか、それが最も重要な問題である。

この問題を考えるために、若干歴史的展望を述べよう。

近代的統計学の源泉を、いずれも19世紀後半に発展したドイツ社会統計学 Sozial Statistik と、イギリス生物測定学 Biometry に求めることは妥当であろう。そして前者はドイツ社会政策学派の思想と、後者はダーウィンの進化論の考えと密接に結びついて展開されたことに注意しよう。20世紀になると、一方ではK.ピアソン、R. A.フィッシャー以降、統計的推測理論の体系が完成し、他方では、農業における実験計画法、工業における品質管理の手法を通じて、大量生産の論理と結びつき、また標本調査法、計量経済学の手法を通じて、政府の政策立案遂行の手段となった。

これらのことは、統計学が後期産業社会（マルクス主義的にいえば独占資本主義段階）の中で生まれ、その論理と密接に結びついて発展したものであることを示唆している。

ここで後期産業社会とは、産業革命によって最初イギリスで生まれた産業社会のより一層発展した形態であり、19世紀後半以降、西欧及び北米で成立し、その後世界各地に及んだものをいう。その特徴は、規格化された大量生産、数万以上の人を抱える巨大企業、官僚化し肥大した国家、都市化した大衆社会、それに対応する大衆文化等によって特徴づけられる社会である。そこで支配するのは「量」の論理である。資本主義的市場経済が支配的であるが、その中で独占力を持つ巨大企業が生まれ、また国家の介入も強まる。ソ連等の「社会主義」も、後期産業社会の一つの変種とみなすことができる。

後期産業社会の基盤を成したのは、フォード・システムに代表される大量生産システム、鉄、電気、機械、化学等の重工業の技術であった。統計学の理論と方法は、このような後期産業社

会の要請に応じて生み出され発展したものであるといえよう。

しかし20世紀最後の四半世紀になって、科学技術の性格は大きく変わりつつある。最近のいわゆる先端科学技術は、マイクロ・エレクトロニクス、バイオ・テクノロジーなど、それまでの科学技術とは基本的に異なった性格を持っている。その変化は俗に「重厚長大」から「軽薄短小」へといわれているが、それに依拠して産業社会の性格も変化しつつある。それは脱産業社会（ポスト・インダストリアル・ソサエティ）ともいわれているが、それが本当に産業社会を脱却したものであるとするのは言い過ぎであるとしても、「大量」と「巨大」を原則とする後期産業社会とは、全く異なった原理を含むものであることは確かである。私としては、むしろそれを高度産業社会と呼ぶことにしたい。

後期産業社会において形成された超大国米ソの対立としての冷戦が、ソ連側の全面的敗北によって終わった後、世界はいよいよ本格的に脱一あるいは高度一産業社会に入りつつあると言って良いように思われる。それは脱規則、自由化、市場競争の強調という点で、むしろ自由主義的資本主義、すなわち前期産業社会への回帰を意味しているかのように見える。しかしそれは正しくない。科学技術の水準も経済の規模も、19世紀前半の前期産業社会とは根本的に異なっている。その上、脱一あるいは高度一産業社会を特徴づける「情報化」という現象は、19世紀前半には全くみられなかったことである。それにもかかわらず、現在の世界がより一層純粋な資本主義の方向に向かっていくように見えますれば、それはやはり過渡的なものであって、またもしそれがそのまま続くとすれば、世界の経済社会は重大な矛盾に達してしまうであろう。

この点についてここで詳しく述べる余地はないが、世界が直面している多くの重大な課題、すなわち人口過剰、貧困、雇用、経済摩擦、民族紛争、資源枯渇、環境破壊などの問題が、資本主義的競争による「市場メカニズムの効率性」のみによって解決できないことは自明のように思われる。それだけでなく、資本主義システム自体の含む矛盾が、供給過剰による大不況、あるいは通貨、金融不安による恐慌という形で爆発する危険も否定できない。

ここでは脱一高度一産業社会への転換がわれわれの統計学にどのような脱皮を迫っているかを考えなければならない。

最初に、現在の統計学の基本的な論理の枠組みが、後期産業社会の原理に対応している点を確認しておこう。蜷川統計学の基本概念である存在たる集団としての「大量」、あるいはフィッシャーの「仮設的無限母集団」、ネイマンの「有限母集団」、あるいはシューハートの「統計的管理状態」のすべてに共通するものは、集団 Mass の概念である。そうしてそれらは同質のもの集まりとして、静態的に捉えられている。そうしてそこから「悉皆調査」、「標本調査」、「サンプリング」などという概念が生まれてくる。それは統計的方法とは、観測を通じて「集団」の状態を把握する手段であるという考え方である。

このような言い方をすると、それは一方では余りにも自明なことをいうと思われ、他方では、いろいろ異なる考え方を何もかも一緒にしてしまったと批判されるかもしれない。しかもやや遠くからみれば、統計的方法が「同質なものの集団」という観念を基礎としたものであり、そのような集団一大量一大衆つまり Mass が、後期産業社会の現実の中に存在するものであったことは、疑問の余地のないところであると思う。

現実に脱一高度一産業社会に入り、情報化が発展する中で、社会に対する基本的情報を提供するものとしての統計や、それを利用するための方法論を与える統計学に対する要望は、当然

に増大すると予想されるが、しかし「統計」という形式での情報が、必ずしも現代社会の要請に適合しないのではないかという印象が持たれることが多いのも事実である。

今年为国勢調査の年であるが、数百億円の費用と莫大な人手をかけた膨大な資料を生み出す国勢調査の結果の中に、果たしてそれだけの費用に値する情報が含まれているか、という疑問は提出できよう。国勢調査によって得られる情報の大部分は、すでに他の調査や統計の中に含まれているし、もし必要ならば特別の調査を別に行うことによって、はるかに低い費用で情報が得られるであろう。さらに最近はかなり早くなったとはいえ、すべての報告書が出終わるまでに数年もかかるようでは現状把握には間に合わないという意見もある。他方、人々の移動性の増大、プライバシー意識の高まり等により、調査の困難は増大している。果たしてそれだけの努力は必要であろうかという疑問もある。

私自身は「国勢調査無用論」は誤りであると思うが、しかし大正9年(1920年)の第一回国勢調査から75年を経て、国勢調査の意味や有用性も大きく変わってきたことは確かである。そこで「何のための国勢調査か」という問いに対して、もう一度真剣に考えてみる必要があると思う。

現在の日本の統計の体系は、戦争直後から構築されたものであり、その大部分は1960年頃までに枠組みが作られている。それは戦後、日本経済が高度成長を遂げて、日本社会が復帰産業社会として大きく発展した過程と並行して発達したものである。しかしその後はほとんど新しい統計は作られていない。そこで現在の統計は、「物不足」の時代に「より多くのモノを効率よく生産し供給する」という目的に役立つように作られたものとなっている。ところが1970年代以降も、オイル・ショック(このときにはエネルギー統計は作られたが)、環境問題、貿易摩擦、円高、バブルの形成と崩壊、企業活動のグローバル化、女性の社会進出、年々多々の問題が生じたにもかかわらず、統計の体系が新たな方向に発展することはなく、むしろ統計調査や統計制度は、行政改革のたびに整理の対象として考えられることが多かった。このことを統計に対する無理解として嘆くことも可能であるが、むしろ世の人々が、「新たな問題に対して新たな統計が必要である」と考えないのはなぜか、ということを検討してみなければならない。

やや古いことになるが、私自身統計審議会の委員やその他の場所での研究会等をメンバーとして、新しい統計、すなわち環境統計、ストック統計等の必要性を強調してきた経験からすると、二つのことが言えると思う。一つには確かに統計情報の有用性と重要性についての一般の認識が不十分であるが、しかしもう一つは、新たな課題に対して適切な情報を得るには、これまでの統計の考え方や制度(いわゆる分散型)の枠組みの中で行われる統計調査は必ずしも適合しない面もあるということである。

現在の脱一高度一産業社会は、多様性と、そして変化の早さによって特徴づけられる。それに対して、不変の集団の存在を前提として、継続性と一様性を重視し、そして各省庁の担当分野に分割して調査を行う現在の統計調査の方向によっては、必要な情報を十分早く獲得することは、極めて困難であることは認めなければならない。

このことは統計調査という概念そのものに反省を迫るものであると思う。と言って、社会情報、特に数値データの獲得、流通について、これまでの統計のあり方から飛躍したものを別個に考えるとすることもできないであろう。統計の基本的枠組みを既成の観念にとらわれることなく再検討し、時代の要請に応じて変えていくことが必要であると思われる。

それに関して若干の点を指摘したい。一つには新しい分野、新しい課題については、「母集団」の設定にこだわることなく、つまりカバレッジや代表性が明確でないデータでも、積極的に取り入れることが必要であるということである。もちろん得られた数字の代表性や信頼性のチェ

ックは常に大切であるが、代表性だけに最初からこだわってはいは、必要な情報が必要なときに得られない。極端な例を挙げれば、大震災が起こった際、ずっと後になってから被害の状況について完全な調査を行うよりも、部分的にでも、なるべく早く、どこでどのような被害が生じているかと報告する方がずっと大切である。この点に関連してまた、後で修正されるような Tentative な値でも積極的に早期に出すべきである。

第二は、継続性を重視しすぎないことである。もちろんデータの定義や性質があまり変わるのも困るが、対象自体が変化しつつあり、新しいものが生まれ古いものが消え去りつつあるとき、データの継続性がなくなるのは、やむを得ないことである。

第三に、本来統計として作られたものでないデータをもっと積極的に利用すべきである。行政や実務の記録から作られる二次統計は、現在では、出生、死亡、貿易など類型化されたものに限られているが、省庁の一般の業務記録や調査結果の中に含まれている情報をもっと活用すべきである。また民間企業や研究機関等の持っているデータも、積極的に利用できるシステムを考えるべきであると思う。

アカデミズムの世界でも、統計学の地位が確立したのは戦後のことであった。

戦前においても、ドイツの社会統計学や K.ピアソン流の統計的方法も若干導入されたが、一つの独立した学問分野としての統計学は、ほとんど成立していなかったと思う。それは大量生産・消費社会としての後期産業社会の形成が日本では遅れていたことと無関係ではない。そうして、日本統計学会や統計数理研究所が戦争直前に設立されたことは、実は全般的軍事化とともに、後期産業社会に入りつつあった日本の現実に対応するものであったといえる（戦後になって、全面的に展開された日本の後期産業社会の基礎が、戦前のいわゆるファシズム期に作られたものであることについては、最近多くの人々が論じているが、注目すべき論点であると思う）。そうして、戦後の日本統計学の確立期に指導的立場にあった人々は、戦争直前期に学問的活動をはじめているのである。

しかも学問としての統計学の全面的な展開は、日本における後期産業社会の発展と同じく戦後のことであった。それは一方では、品質管理、農事試験、標本調査、計量経済モデル分析、薬効試験等の現実問題と結びつき、他方では戦後民主主義の思潮に乗って、特に初期には華々しく開花した。最初は「推計学」などと呼ばれた当時の「新しい統計学」は、既成のアカデミズムに対して若干の距離を置く場所にあり、そのことがまた一つの魅力であった時代でもあった。

しかし多くの人々の努力によって、統計学の T. クーン流に言えば、「通常科学」Normal Science としての立場は確立された。大学院における専門研究者の養成も、小規模ながら確実に行われるようになった。

しかし最近になって、統計学は世界的にいわば成熟段階に達してしまったように思われる。行き詰まりといっは言い過ぎであろうが、少なくともこれまでの理論の枠組みの中では、研究の努力に対する成果の限界生産性が低下する段階に達したことは確かであると思われる。応用分野は特に医療分野などでは、大量生産工程における統計的管理状態の達成と言うような古典的な考え方は、もはや有効性を失っているのではないかと思われる。

他方コンピュータの発達は、単に計算や情報処理技術の体系にとどまらない情報科学、コンピュータ・サイエンスの発展をもたらした。それは、それ自体豊かな内容を持つと同時に、心理学、言語学者等の人文科学分野を含む、多くの科学分野に影響を及ぼしつつある。統計学においても、複雑な計算手法を極めて簡単に適用できるようになったこと、結果のビジュアルな表示が容易になったこと等によって、既にかなりの影響を受けているが、それはまだ、これまでのところでは計算技法の上での表面的なものにとどまっていると言わねばならない。

しかしもう少し立ち入って考えると、統計的データを観測し分析するという過程そのものを、一つの情報過程として捉え、統計学を広い意味の情報科学の一部門として位置づけることが必要ではないかと思われる。そこでこれまでの統計的方法の理論的枠組みは、変わらざるを得ない場合があるように思われる。

古典的な統計的方法の論理的順序は、
対象分野についての概念的枠組みの設定

↓

それについての基本的な理論モデルの想定

↓

関連するデータの収集または調査、実験等によるデータの獲得

↓

データに関する統計的モデルの設定

↓

統計的方法による結論の導出

↓

対象に対する理論の検証、具体化、あるいは具体的行動の採用
と言う形になる。この一部はフィードバックして繰り返されることがある。

しかし現実の情報過程においては、具体的な行動とデータ獲得の過程とは、切り離し難く結びついている場合が少なくない。医学の臨床の場合には、患者の治療と観察とは、原理的には切り離すことができず、治療しつつ観察し、その過程の中で次の行動を決定することにならざるを得ない。この中で統計的方法の論理に忠実な二重盲検法などが適用できる場合は限定されているように思われる。それ以外の場合から得られる情報も多いはずである。

私はここで二重盲検法を否定するつもりはない。私が強調したいのは、そのような「統計的に正しい」方法が適用できない場合、あるいは適用されなかったデータに含まれている情報、とくにいわゆる「カルテ」に含まれている膨大な情報を、どのように利用するかということなのである。現実の臨床家は、もちろんこのようなデータを利用して「経験」を積んでいるのであって、それをどのように客観的なものとするかは、統計学にとって重要な課題であると思う。

同様な課題は他の面でもあると思う。例えば、最近政府の景気の局面判断が、現実に対して遅れ、そのために適切な政策が発動できないことがあるといわれる。その原因の一つは景気判断のための統計資料が早く得られないという点にあると思われる。このような場合、国民経済計算などのいわば「正式の統計」が得られる前に、カバレッジ、信頼性はそれほど高くなくても、景気の動向に敏感な指標に基づいて早く判断を下すことが必要であると思う。そのような指標を構成することも一つの課題である。

ただしここで見逃してはならない一つの点は、統計学がデータを分析する方法と同時に、データのあるべき姿、つまりデータの取り方についての規範を与えるということである。ランダム化などはその点で評価しなければならない。データの批判的吟味ということは常に重要である。

近代統計学も100年以上の歴史を経て、後期産業社会から次の段階への歴史の転機に当たって、統計学がこれまでの成果を生かしつつ、従来の枠組みにこだわらず、新たな展開と発展を遂げるべき時期に来ていると思う。

わが日本統計学会の若い世代の人々が、大胆で新鮮な研究を深めて、このような発展に貢献されることを願いたい。