

先端技術社会における統計学の役割†

竹内 啓*

The Role of Statistics in Recent Technology Oriented Society

Kei Takeuchi*

「統計」は先端技術社会の数値的基礎を与える。その役割を考えるときは、「統計」と「統計的データ」を区別した方がよい。統計的データは実験や調査で得られるデータで、統計的方法の解析対象である。統計的方法の役割については、タグチメソッドの有効性で例示されるような発展方向に注目しなければならない。統計は、社会集団を論量的に把握したものである。社会構造が複雑になっている現代では、統計の対象となる集団概念のダイナミックな変化にも注意しなければならない。統計は公共性を持つから、その公開と提供のサービスを怠ってはならない。統計を社会に供給する形態としては、コンパクトな要約が有効である。それと同時に、個票に含まれる情報をプライバシーを侵さない範囲でなるべく多く提供することも重要であり、そのための工夫が必要である。

1. 先端技術社会における数値的基礎情報

先端技術社会は、先端科学技術が大きく発展し、普及する社会である。先端科学技術の中で、マイクロエレクトロニクス技術を基礎とするコンピュータ技術、情報通信技術が重要な意味を持つことはいうまでもない。先端技術社会は、これらの技術による大量の情報処理、情報伝達、情報蓄積が行われ、各分野で利用される、高度情報化社会である。先端技術社会と、高度情報化社会は、概念的には同じことを意味するものではないが、先端技術社会が高度情報化社会という一面を持つことは確かである。

統計は社会全体に関する数値的基礎情報を与えるものとして、高度情報化社会においては原理的にその重要性をますます増大させるはずである。またコンピュータ、情報通信技術の発達は、統計情報の高度加工と結果の適時利用とを可能にするはずである。

2. 統計と統計的データ

実際コンピュータの発達は、高度複雑な計算手法の利用を容易にし、普及させた。僅か30年ほど前には、分散の計算さえ容易でなく、数個以上の説明変数を持つ線形回帰式の係数を最小2乗法によって計算することは、極めて困難であった。

現在多変量解析のいくつかの手法、主成分分析法、因子分析法、数量化などが、多くの分野で広く応用されているが、このことはコンピュータの発達と、ソフトウェアパッケージの普及なしには考えられないことである。

ここで固有の意味での「統計」と「統計的データ」の区別に注意しておかねばならない。前者は数値的情報を得ることを目的とし、主として社会的集団を対象とした統計調査によって得

* 東京大学先端科学技術センター, 〒153 東京都目黒区駒場 4-6-1

† : 本論文は、日本統計学会60周年記念事業の一環として、1991年11月15日東京で行われた、シンポジウムで発表したものをまとめたものである。

られたものをいい、後者はより広く、観測、調査、実験等によって得られ、統計的方法による分析の対象となるデータをさす。

統計的データを利用し、統計的方法を応用する分野が、しだいに拡大し、またその中心が移って来たことは事実である。最近では、医学、薬学等の分野への応用が盛んになっている。

3. 統計的方法の役割

しかし先端技術社会における統計的方法の利用、普及は決して直線的に拡大してはいないといわざるを得ない。むしろ先端技術社会は統計、統計データ、統計的方法のあり方に対しても、再検討の必要性を生み出しているのである。

現在先端技術の多くの分野で、日本産業が極めて強い競争力を持ち、そうしてその競争力の源泉が日本産業の製品の高い品質と信頼性にあることはよく知られている。そうして日本の製品の品質向上に統計的品質管理が大きく貢献したことも広く評価されている。

しかし固有の意味の「統計的」品質管理の方法の重要性は、先端技術の中ではむしろ少なくなっているといわねばならない。しばらく前までの「重厚長大」型の大量生産工程においては、大量の製品の集団の中の「不良率」や「ばらつき」が問題となり、それを統計的な方法によって処理し、管理することが可能であった。しかしマイクロエレクトロニクス技術や、バイオ技術の最先端部分では、統計的に処理できるような「不良率」や「ばらつき」の存在は許されない場合がふつうである。また先端的技術の分野では「最適水準」を統計的実験計画法によって求めるといふこともほとんどない。

勿論先端技術の時代になっても、実際に製品化し、生産工程を安定化する段階では、統計的品質管理、或は統計的実験計画の方法が依然として有効である場合は少なくないと思われる。特に最近では、製品の物理的品質ではなく、それが現実に用いられる場における機能のばらつきを制御することを目標とする田口玄一氏の開発した方法、いわゆる「タグチメソッド」が国際的にも注目されている。それは工業における統計的実験計画法の一つの発展形態と考えられよう。

バイオ技術に関連した、医学、薬学、農学、食品工業等の分野においても、実際の応用の場では統計的方法が用いられる場合は少なくない。

これらの分野における統計的方法の発展についても論ずべきことはいろいろあるが、ここでは立ち入らないことにしよう。

4. 統計情報の提供の形式

ここで主として問題にしたいのは、固有の意味における「統計」、つまり統計調査の結果としての統計データの、先端技術社会における意義と役割である。

私が統計審議会委員として、情報処理部会長を勤めていた頃、民間から官庁統計の結果表を磁気テープの形で提供してほしい、またそれを加工して再提供することを可能にしてほしいという要望がしばしば寄せられた。このような要請に対して、官庁側でも対応し、多くの省庁でテープ提供、再利用許可をする体制が作られた。ところがその後の実績を見ると、現実に磁気テープによる提供が求められた件数は極めて少なく、いわば、期待はずれであったといわざるを得ない。

また統計データを収集、蓄積した統計データベースにしても、商用のものとしてある程度広く利用されているのは「日経 NEEDS」だけであり、商用でない、官庁、地方公共団体、あるいは民間法人等におけるデータベースにしても、業務用データベースでない、一般的な統計データベースは、構築されることがあっても、十分に利用されることはあまりないようである。

このことは高度情報化社会になっても、少なくとも現在の形での統計情報データベースに対する需要はあまり大きくないことを意味しているといわざるを得ない。

それには二つの面がある。一つは現在でも統計データを用いてマクロ経済的な大体の動向を見るには、「統計要覧」的な一冊の統計集で十分であり、磁気テープによる収録を必要とするような詳細な統計表は不要な場合が多いということである。そうして統計数字を利用する場合のほとんどが、このようなマクロな数字の利用であると考えられ、このことは今後も変わらないと思われる。

問題に第二の面は、逆に詳細な情報を得ようとする、最も細かい結果表でも十分でないということである。個々の企業や個人についての情報を求めることは、統計調査の主旨或は統計法の精神からして排除されるとしても、詳細な地域別、産業別等のデータ、或はいくつかの量や特性の間の関連、同一の対象における時間的变化、などを知りたい場合、最も細かいもので2～3の基準によるクロス分類表、そしてその中の各セルにおける対象の個体数や平均値が表示されているだけでは情報として不十分であり、また標本調査の場合、精度も足りない場合が少なくない。個々の対象についての情報を求めるのではない場合でも、個票にさかのぼらなければ必要な情報が得られない場合は少なくないが、個票を利用することは厳しく制限されていることは周知の通りである。

統計調査における秘密保持、或は個人のプライバシー保護（この2つは本来同じことではない）の観点から、個票利用に制限があるのは当然であるが、現状における統計情報の発表形式とその利用に関するルールには硬直的すぎる面があるので、もっといろいろな工夫がなされ、またそのような工夫が現実に生かされるよう、制度的な改善もなされることが望ましい。

5. 伝統的な統計概念

しかしより根本的に、統計によって先端技術社会に関する適切に有用な情報が得られるようにするには「統計」というものの考え方の基礎に、さかのぼって再検討すべきところが少なくないと思われる。

伝統的な統計学の考え方によれば「統計」とは一定の社会集団を対象として、その構成単位の特定の特性を調査し、集団の中でのその状況を数値的に表現したものであると考えられる。その場合、実際に調査される対象は標本調査の場合のように集団の一部であっても、関心はあくまで全集団の状況にあり、個々の単位にはない。そうして集団を構成する単位は基本的に同質なものと考えられ、全体について一定の方式で一様に調査することが基本的と考えられている。

統計についての基本的概念にもとづいて、統計調査の手続きが形成されて来た現在では多くの点でそれは明確に確立されている。それは貴重なものであるが、しかし現代では伝統的な形式に過度にこだわることは、統計情報の有用性を損うことになりかねないと思う。

6. 集団という概念の動的な理解

基本的な問題は、伝統的な統計概念が本質的にスタティックであって、変化の激しい現代には適合し難い面があるということである。勿論時間的な変化を追う「動態統計」の概念は、伝統的な統計学の中にも重要な位置を占めている。しかしその場合でも対象となる社会集団自体は不変なものにとどまると想定されている。ただ集団全体における特性値の傾向は時間的に変化するもので、それを明らかにするのが動態統計であると考えられている。

しかし例えば現代の産業統計において、一定の産業部門における「企業集団」が、安定的に存在し、それを調査することによって特定の産業の状況を把握することができると考えたので

は、恐らく有用な情報は得られないであろう。現在多くの企業は、戦略的にその経営内容を多角化、多部門化し、またそれを状況に応じて変化させている。更に「分社化」によって形式上は独立で、しかし事実上は本社の一事業部門或はそれ以下にすぎない、多数の会社を作り出している。このような場合には企業集団の把握について、これまでのものとは違った考え方が必要になる。

先端技術社会の特質が多様性と変化の速さにあるとすれば、統計がスタティックな概念規定にこだわっているのは、時代の変化に対応できないということになるであろう。

7. 完全性と柔軟性

統計においては、完全性、正確性、継続性が重要視される。勿論これらのことは大切であるが、しかし現代ではこれに加えて、適切性、適時性、柔軟性も要求される。つまり全体についての完全なデータでなくても重要な意味をもつところについての情報が、少々の誤差はあってもなるべく速く、そして社会の変化に対応してより適切な形で、提供される必要がある。

このような場合には、完全、正確、そして継続的な、いわば統計表「原本」のほかに、部分的、暫定的、そして試験的な性格を持った情報を、その性格を明確にした上で提供することが望ましい。一部の統計については「確定値」の前に「速報値」が計算され公表されているが、このようなことはもっといろいろな形でなされるとともに、その精度を上げるための工夫も必要とされよう。

これに関連して統計分類の問題がある。私は統計審議会分類部会長として、標準分類(産業分類、職業分類、商品分類)の改訂にたづさわった経験から、分類を社会の変化に対応して変えて行く必要と、データの経時的比較を可能にするために継続性を維持しなければならないこととの、バランスを取ることの難しさを痛感した。標準分類がある程度「保守的」になることは止むを得ない面もあるので、むしろ今後は副次的分類、クロス分類等を標準分類と併用して行くことが望ましいであろう。

8. 情報公開と情報サービス

もう一つの基本的な問題は、統計が誰によって誰のために作られるかという点である。伝統的な考え方によれば、統計は国家或は政府が、統治ないし行政目的のために作成するものであるとされていた。統計が公表される場合でも、国民に情報を提供するという目的は副次的と考えられていた。

先端技術社会が高度民主主義社会でなければならないことは、最近の官僚専制的な社会主義体制の解体によってますます明らかになって来たと思う。その中で民主国家における官庁統計の意義と役割についても、再考する必要があるだろう。そこに二つの面がある。

一つは「情報公開」の面であり、政府が政策を立案し、施行する中で、その根拠とする事実、およびそれに関する情報を国民の前に明らかにするということである。それは民主主義社会では、国民に政府の政策を批判するための情報を与えるという面も持つ、政府に都合のよいように統計が歪められたり、情報が秘匿されたりしてはならないという原則は以上のことから生ずる。

第二は「情報サービス」の面である。すなわち国民、或は企業ないし個人が、いろいろな形で必要とする社会的情報を、国民に提供するということである。或は社会的情報インフラストラクチャーを提供するといってもよい。大規模な統計調査を行うことは、民間では資金や人手の点からも、場合によっては必要とされる信用と権威の点からも、不可能であるから、政府がそれを行うことは「公共サービス」の一つと考えられる。

こういう点からは、官庁統計の提供に関しても、単に行政目的で作られた統計を「民間にも」、利用させるというのではなく、最初から国民の要請に応じた統計データを作ることも考えられてよい。

統計の「公共性」という観点は、民間機間の行う統計調査にも適用されるべきである。すなわちプライバシーの保護、結果の公開等について政府統計と同様の基準が守られるべきである。そうして民間統計が政府統計を補い、或はそれに代るものとして利用され得る部分も少なくなないので、「統計」体験を全体として考える時、民間の統計調査もその一つの構成要素と考えなければならない。

先端技術社会における統計の問題を考えると、重要なことは情報処理技術や、統計的計算手法ではなく「統計」概念そのものの吟味であることを強調しておきたい。