

て、日本経済統計文献センターにおいて、今後開発すべきデータベースのあり方を定める基礎資料にしたいというのが、これらの文献が作成された意図なのであった。統計データいじりを職人としてやってきた私は、コンピュータには弱く、データベースを論ずるには不適任である。だから、以下においては、主として統計の現状と歴史統計をめぐる論議について感想を述べ、データベースについては簡単にふれることで責を果させていただきますと思う。

現在の統計について。日本の統計は、国際的にみても、もっともよく整備されている部類に属する。しかし、そのなかに多くの問題が秘められていることも、また否定できない。問題点としては、次の諸点が指摘されよう。

① 各種統計の整合性のチェック、精度の検討等、資料批判。その点についての学会や雑誌による公開、さらには作成者と利用者の間での討議の場の恒常的な設定。② 統計データの所在情報とその使用可能性の確保。③ 個票その他マイクロ・データの保存と利用可能性。

整合性の問題は、さらに、a. 事業所統計、工業統計、法人企業統計など、事業所、企業ベースの構造統計相互間のミス・マッチ、b. 時系列標本調査の標本切替時点における不連続性の問題などに分割される。この両者のうち、前者については松田教授のくわしい検討がなされている。すなわち工業統計と事業所統計の間には資本金規模、従業員規模別の企業数について、系統的な不一致がみられる。たとえば従業員規模別にみれば、従業員499人以下の事業所数は、工業統計の方がつねに多く、500~4,999人のそれは事業所統計の方が多い。また、法人企業統計と工業統計の間では、資本金規模1億円以下の企業数はつねに法人統計の方が多く、1億円以上では逆転する。また付加価値額も、法人統計の方がつねに多い。松田教授のいうように「企業と実際の製造に関する諸活動がうまくリンクされた形でデータが出されていない」のである。この点について、工業統計における企業編と産業編のデータを使い易いように再集計し、そのファイルを共用化できないか、という提案がなされている。

構造統計の多くが、相互に整合的でないことは少し分析を行なった人なら、ほとんど気づいているであろう。税務統計や商業統計などをとりあげれば、不整合の程度はもっと著しくなるかもしれない。シンポジウムに紙上参加するつもりで私見をつけ加えれば、その理由として① 調査の方法、② 個票の精度、③ 調査対象の把握率、④ 審査のやりかた、⑤ 分類格付のしかた、など

尾高煌之助・松田芳郎編

『日本経済統計データベース編成の 課題と方法』

——シンポジウムの記録——

一橋大学経済研究所 1983.3 x+261ページ
日本経済統計文献センター

本報告書についての所感の執筆を引き受ける位なら、熟があってもシンポジウム(私もお招きを受けていた)に出ておいた方が何倍か楽であったと後悔している。この本は、歴史統計と現代統計についてのあらゆる問題点の縮図であり、また一橋大学附置日本経済統計文献センターのデータベース作成についての、多種多様な構想と要望の集成であって、書評の対象としては、もっともやりにくい部類に属する。紹介することだけでもほとんど不可能である。その難題に取り組まざるを得なくなったのは、評者の迂濶のゆえであり、また統計マン同士の義理人情の結果でもある。

愚痴をさておいて本題にはいろいろ。本書は松田芳郎教授による総合報告と、東京と京都における2回づつのシンポジウムの記録とから成る。シンポジウムは、両地とも、現代の統計をめぐるものと歴史統計をめぐるものから成っている。また、シンポジウム実施の後昨昭和58年に行なわれた『日本の社会経済統計データベース需要動向調査結果報告書』*は、経済学・社会学研究者のうち日本についての実証研究を行なっていると思われる2,213名について、統計データの利用の実態について、くわしい調査を行なったものである。この両者をあわせ

がどうなされたかを考えなくてはならない。それらはいまって結果のひずみをもたらすからである。統計資料を批判してゆく場合、この5つのいちいちを吟味することが、やはり必要であろう。市野・腰原両氏の発言(65~66ページ)や、のちの田中・市野・柳川氏らの発言(87~97ページ)は、重要な点をふくんでいると思う。一次統計の作成のさいの問題が含蓄をこめて指摘されているからである。いいにくいことだが、気づいたいくつかの点を指摘しておくことにしよう。まず、調査方法の点で、センサスないし準センサスの場合でも、また標本調査の場合でも、対象のカバレッジは決して高いものではない。おそらく税務統計のようなデータを母集団に利用できないかぎり、かなりの調査洩れが伴うことはさげがたい。また個票の精度は決して高いものではない。所得や売上高や在庫高などを良心的に答えるものは必ずしも多くはないのである。統計作成の現場にいる人たちは知っているが、利用者にはあまり知られていないこの点を考えるとき、各種統計相互間の不整合性の問題をどこまで精しく論じうるかという問題には、限界があることを考えざるをえない。世帯の所得についての2つの調査が同一世帯についてくり返し行なわれたとき、両者の結果の相関係数は0.8ていどにすぎなかった実例さえあるのだから。分類や格付も、実際には問題である。とくに現在のように企業の事業内容が多面化しているため、製造業かサービス業かという大分類の上での移動が生じる可能性さえあるわけだから、ここにもミス・マッチの原因がひそんでいそうである。

この種の問題があるにもかかわらず、私は統計分析一般について懐疑的ではない。むしろ、古典的な大数の法則は今も機能していると考えている。しかし、現行統計の結果のギャップを、ある程度以上に縮小することは困難と、諦めを付けているのが私の本音なのである。そのことは個票利用の問題にもつながる。個票の利用は、現行統計法の制約もあり、しかもプライバシー保護の要求が強まっている状況下では容易に進展しにくい問題のように思われる。むしろ、現在の文書による公表以上に詳細な多元集計データをプライバシーの問題が生じない範囲で作成し、その情報を磁気テープなどの形で利用可能にすることが、当面望みうる妥当な解決策なのではないか。それはまた、データベースのあり方にも、関係するところが大きいかもしれない。もっと大胆に個票データの公開をみとめるように統計法を改正することが、研究者の立場からは望ましいとしても、調査対象の立場との調和が必要であって、統計利用者ないし分析者の立

場のみを主張しても無理であろう。そこに現代の統計をめぐる最大のジレンマの1つがある。

これに比べれば、時系列データの問題は、むしろ母集団データの迅速適確な整備が行なわれさえすれば、かなりの程度解決可能な問題なのかもしれない。「毎月勤労統計」にしても、「法人企業統計」にしても、税務統計の母集団を使うことができるようになるか、事業所統計の母集団の半年ないし1年ごとの中間的整備が可能になりさえすれば、抽出替えを頻繁に行なうか、ローテーション・システムをとり入れることによって、バイアスを小さくすることができるはずだからである。だが、それと同時に、本書の中でも松田教授によって指摘されているように、標本のローテーションを行なうなどの調査方法を採用することなく、伝統的な調査をくり返してきた調査主体に対しても、批判は向けられるべきであろう。新しい手法の採用や、新しい統計の開発について、1950年代まで活発だった官庁の統計担当部局は、現在においてはいかにも消極的の感が深い。定員・予算の制約、地方の統計部局の反対、調査員の確保の困難、調査対象の非協力などの困難が山積しているとはいえ、このままでは、アップ・トゥ・デートな統計需要にこたえることは難しくなりつつあるのではない。

歴史統計について。松田教授の画期的業績として、明治期の会社統計を、会社興信録などの資料によって、大幅に拡充改善されたことがある。この種の仕事は、明治期の経済についての通念を一変させるほどの意味をもっていると私は考えている。松田教授にとっては、これはデータベース拡充の実験にすぎないのかもしれないが、お世辞ではなく、ご本人が気づかない衝撃を秘めていることを指摘しておきたい。

さて、歴史統計にとって最大の急務は、やはり埋もれた資料の発掘、整理にあると考える。黒崎教授や今津教授が指摘されるように(115~116, 123, 195~196ページ)、警察統計とか、衛生年報とか、学事統計とか、博覧会資料とか、見落されている基本資料は意外に多いのである。いちばん地味で手間のかかる仕事ではあるが、それらを1つの機関が蒐集、保存して下さることを、心から期待する。そうした基礎の上に、データベース化が積み重ねられるべきであろう。

歴史統計において、さらに問題とすべきことは、その時系列的連続性や、地域間の比較可能性である。統計表章上の定義、概念等はしばしば変更され、その注記も付せられていないデータが多い。たとえば税制の変更などととも、数字が突然変化する場合がしばしばある。こ

うした問題について、納得のゆく説明が行なわれることは、歴史統計の利用上不可欠の配慮であろう。このように不安定なデータを利用して推計を行ない、加工統計を作成することは、当然ながら困難であり、誤差と偏りを伴うことは避けられない。その意味で加工は個々の研究者の責任においてなされるべきものであり、センターがそれに乗りだすことにはおのずから限界があるろう。ついでながら、すべてのデータを、SNA との関連で整理してゆこうとする国連の発想には、私は賛成しがたい。その時点の統計作成の目的や方法が、SNA を予想していなかったのだから、その事情を忘れるわけにはゆかないであろう。松田教授の生データ主義に私も賛成したい。

なお、一橋大学のセンターの所蔵資料目録の発刊について、松田教授はむしろユニオン・カタログが望ましいとして、消極的であると述べられている(191 ページ)。しかし、理想はとにかく、成立以来20年弱の蓄積を学界の共有財産にするために、やはり発刊にふみ切っていたことを私は強く希望したい。ローマは一日にして成らず、である。

データベースの問題について。すべてのことを、センターに期待するのは、人手と予算の制約がある以上、無理であろう。また松田教授をはじめスタッフ諸氏の志向がまず尊重されるべきことはいうまでもない。統計の利用者はもちろん、データベースの作成者をも含めて、す

べての研究者は、多かれ少なかれ我儘で自分中心主義なのだから。そのことを認めたくえで、希望を出すとするれば、次の諸点があげられる。ひとつには、LTES の基礎となった、もしくはその改善の基礎となるべき歴史統計のデータベースの作成に力をそそぐこと。ふたつには、実験的に、工業統計や事業所統計についての細部のチェックを実験的に行なうこと。そして、一般データについては、むしろ市販のデータベース等にゆだねて問題にしないこと。

なお、研究者の統計需要調査は、それ自体として極めて興味ふかいものであった。なかでも、公刊統計データの存在が意外なほど一般に知られていないことは、私には驚異であった。その入手が困難であることもこれほどとは思わなかった。この点は、統計作成者である諸省庁に、周知のための配慮を要請しなくてはならないであろう。たしかに、現在の諸統計は、行政上の目的を優先して、作成され、配布されていることは、否定しがたいが、同時に、研究者の利用によって、はじめてその価値がたしかめられることもまた事実なのである。〔中村隆英〕

* 松田芳郎編『日本の社会経済統計データベース 需要動向調査結果報告書(概要編)』一橋大学経済研究所日本経済統計文献センター、1984(統計資料シリーズ No. 28), x+155 pp.

農業経済研究 第56巻 第1号

(発売中)

《論文》

沼田 誠: 大正・昭和期の茶業地帯における農家経済の展開構造
——静岡県土方村の事例より——

津田 渉: 十勝畑作地帯における中上層農の展開と生産組織化の動向
——芽室町の実態——

藤田夏樹: フロー構造の定量化について
——尾崎巖氏の単位構造の一般化——

《研究ノート》

玉真之介: 東浦庄治の日本農業論

《書評》

山内豊二著『農業災害と農業保険の国際比較——先進国と開発途上国の対比において』(家永泰光)

小林康平著『牛乳の価格と需給調整』(永木正和)

川野重任編『技術移転と文化摩擦』(逸見謙三)

川本 彰著『むらの領域と農業』(渡辺兵力)

B5判・60頁・定価1200円 日本農業経済学会編集・発行/岩波書店発売