

# マイクロデータの利用特性と統計利用論

森 博美（法政大学・経済学部）

## はじめに

統計は、その作成者（統計作成機関）がそれまで把握していなかった個体に関する情報を何らかの方法で獲得し、その結果を数値として表記したものである。統計作成者が統計作成において獲得の対象とするのは、個人、世帯、企業といった調査対象に関わる属性やそれらの活動実態等に関わる情報であり、以下ではこれを「統計原単位情報」と呼ぶことにする。調査実施者が調査客体から獲得する統計原単位情報には、統計作成に使用される項目（統計項目）だけでなく、統計作成の対象となる個体（調査客体あるいは被調査者）の識別情報である世帯番号、企業・事業所番号、氏名（企業等の名称）や住所（所在地）、電話番号等といった個体識別情報も含まれる。個体識別情報については、回収された記入済み調査票についての照会の際の連絡先としてまた他の統計との連結（リンク）の際の連結情報という形で統計作成に関わってはいるが、それら自体が直接統計に編成されるわけではない。なお以下では、個体識別情報を含めたより広い概念を広義の統計原単位情報とし、また特に誤解が生じない場合にはそれを除いた統計に編成される個体情報のことを統計原単位情報と便宜的に呼ぶことにする。

統計の作成方法には、統計原単位情報の獲得方式の違いによって次の2つがある。統計作成を自己目的として独自に設計した様式（調査票）への回答（報告）を調査客体に対して求め、得られた統計原単位情報を編集することで統計を作成する方法と、統計とは異なる目的のために収集された情報の一部を統計原単位情報として統計作成に転用することで統計が作られるのがそれである。前者のような過程を経て作成される統計は一般に「調査統計」、そして後者は「業務統計」と呼ばれている。

本稿では、政府が統計調査によって作成する政府調査統計について、特に統計原単位情報の獲得形態ならびに結果資料の提供形態という2つの側面からわが国における政府調査統計の展開過程を跡づけるとともに、マイクロデータがこれまで統計作成・提供の基本的在り方とされてきた集計表によるデータ提供と統計利用面でどのような違いがあるか、また内外におけるマイクロデータ提供の現状、さらにはわが国における統計の利活用促進に向けての課題などについて以下に検討してみたい。

## I. 統計の調査形態と結果資料の提供形態

### 1. 近代統計形成期の政府調査統計の成立

#### 1) 表式調査による統計原単位情報の獲得

個々の統計調査の開始には、その時々の政策課題や海外における統計整備の動向といったような時代の要請が深く関わっている。わが国の近代統計形成期における政府統計調査は、いったいどのような時代的要請を契機に作成が開始され、そのことが統計調査の形態にどのような特徴を付与しているのだろうか。わが国における近代政府統計の黎明期にあたる明治初期に作成開始された生産と人口分野でのいくつかの政府調査統計についてその特徴をまず検討してみよう。

租税収入の8割以上を地租に依存していた明治新政府にとって、国家歳入の確保は最優先の政策課題であった〔(2)p.5〕。維新の翌年に会計官（その後大蔵省と改称）は、「村高取調」に付帯させる形で「物産ノ事 但村々ニテ一箇年分目当高ヲ管轄中惣括ニシテ申立候事」（第398号：〔(1)p.43〕）として物産報告を命じている。これは、翌1870年（明治3年）に、民部省地理司による物産表の報告提出命令第623号へと引き継がれる。「物産表」調査の始まりである。「物産表」は、間接税制整備のための物資流通量の把握を目的に、農水産物や工業製品<sup>(注1)</sup>の生産数量の報告を各村々に求めたものである。1872年（明治5年）には、大蔵省通達第37号により郡別に報告をするように改められた。

この「物産表」は、一定の調査様式に生産者毎の生産量や収穫高を記載するのではなく、村や郡計を所定の欄に記入する「表式調査」方式で調査が実施された。報告者が記入する統計原単位情報が個体情報ではなく集計量である点がこの調査形態の最大の特徴である。

「物産表」がこのような表式調査の形態をとった理由としては、報告結果の集計能力が十分ではない中で、ともかく生産総量の迅速な把握を行うことに調査実施者の目的があったことによるものと考えられる。また、「物産表」が基本的に物資流通量の調査という性格を持つことから、各物資の生産条件や生産構造に関わる事項は、調査の埒外に置かれていた。

通達に明記されていたのは報告すべき物産名のみで様式自体の作成は各県に委ねられたことから、報告される計量単位が地域毎にまちまちで、統一性に欠けるものであった<sup>(注2)</sup>。さらに深刻な問題として、ストック量である現在高とフロー量である年間生産量という異種の統計数値が混在しており、結局、全国総計の算出もできなかった。従って、「物産表」には個体の個々の属性に関する報告項目はなく、生産条件や生産構造を知る手がかりとなる各種属性別の集計表は作成することができない。

税制の整備とともに、教育、徴兵、郵便等の制度化を急ぐ政府にとって、人口統計の整備もまた焦眉の課題であった。そこで政府では、1871年（明治4年）に治安維持的性格も併せ持つ「戸籍法」を布告し、それまでの「人別改」に代わる人口把握の体制を整えるこ

とになった。翌 1872 年に、「戸口調査」による戸籍登録はこのような政策的意図の下に実施されたものである。

「壬申戸籍」と呼ばれるこの人口登録は、1920 年（大正 9 年）の第一回「国勢調査」までの唯一の全国ベースの静態人口データであった。それはまた、人口推計のベンチマークの役割も果たした。なお戸籍表と職分表から構成されるこの「戸口調査」は人口統計の一つの出発点ではあったが、欧米の主要国ですでに実施されていた人口センサスとは多くの点で異なる性格のものであった。

「物産表」や「戸口調査」の他にも、明治新政府は、1873 年に地租改正のために全国の地価の評価と所有者の登録を行う目的で「全国土地調査」を実施している。このように、明治初期に実施された政府統計調査では、統計作成は徴税や治安維持さらには殖産興業や救貧といった行政目的に直結したものとして開始されたものが多く、調査統計でありながら、行政資料的要素の強いものであった。また統計原単位情報の収集も、調査員による被調査者（報告客体）からの直接的徴収ではなく、府県や郡、村といった行政組織を通じた表式調査による徴収が支配的であった。

このように、わが国の初期の政府調査統計は、全国統一的な調査様式の導入、調査事項に関する概念や記入単位の統一、調査結果の徴税その他の行政目的への使用禁止、総計から属性別集計への移行、などを通じて直接的行政目的での情報徴収から次第に自立した統計作成そのものを目的とする調査へと脱皮していくが、中でもその最大の契機が個票の導入である。

## 2) 政府調査統計の新たな展開—個票による統計原単位情報の獲得

「物産表」として開始されたわが国の産業統計、特に工業統計は、殖産興業政策の下で発展を遂げる。1894 年に「農商務統計報告規程」により工場統計にはじめて個票が導入され新たな段階に入る。しかし農業その他の分野では、なお引き続き表式調査方式で統計原単位情報の収集が行われた。農業分野での最初の個票による調査は、農林省による 1938 年の「全国農家一斉調査」まで待たねばならず、農業分野で表式調査に終止符を打つのは、「農事調査」を農林省統計に統合した 1940 年の「農林水産業調査規則」によってである。

一方、人口統計はやや事情が異なる。明治初期の政府の通達による人口調査はあくまでも戸籍を中心としたもので、そこには人口の年齢構成や配偶関係といった自然属性あるいは移動や職産といった社会属性を把握するといった発想は見られなかった。

この時期の調査で異彩を放つものとして、1879 年に太政官政表課が実施した「甲斐国現在人別調」がある。この調査では早くも個票が導入され、職産分類では「戸口調査」より明確な職業、産業上の位置づけがなされていた。明治期の指導的統計家杉享二によって企画立案されたこの調査は、欧州における人口学に理論的基礎を置く近代的な人口センサスのモデルとして、近く実施される予定の「国勢調査」の試験調査的性格を持つものであった。このため、当初、他の府県でも実施する最初の調査として企画されたものであるが、それは実現せず、結局、わが国全土と対象とした個票による政府の人口調査は、大正 9 年

(1920年)の第一回「国勢調査」まで待たねばならなかった。

このように、政府調査統計への個票の導入は、統計作成が行政目的での情報徴収から自立していく中で、単に総人口や生産・販売総額だけでなく、その構造や属性別特徴の解明を統計作成の重要な課題として設定する西欧の近代統計調査のわが国への導入と密接に結びついている。

## 2. 表式調査から個票調査へ

### 1) 表式調査における集計制約

初期の「物産表」では、報告すべき品目の名称だけを指定する形で報告が徴収された。このため、各府県では独自の様式に基づき、郡・村からの統計原単位情報の徴収を行った。その結果、報告事項についての計測単位や把握の形式が不統一で、収集された資料は全国総計の算出もできないものであった。しかしこのような初期の報告徴収に特徴的な収集資料の欠陥の責めは表式調査そのものではなく、その実施体制の未整備に起因するものである。なぜなら、全国统一の報告様式と静態量、動態量の区別も含め報告にあたっての統一された計測方法(統計基準)さえ徹底されれば、こういった問題は解消されるからである。その意味では、表式調査という調査形態でも、個票調査と同様、統計対象としての静態量、動態量のそれぞれについて統計原単位情報を収集することができる。むしろ両者の違いは、集計過程にある。

表式調査の場合、報告原票となる様式それ自体に集計量が記載される。表式調査は、それを積み上げることで、村計、郡計、府県計そして最終的には全国計を容易に求めることができる。手集計という計算技術面での制約の中で短時間のうちに全国計を算出する最も合理的な統計原単位情報の収集形態であった。

表式調査で、生産量や生産額に加えて報告対象地域に所在する農家数や工場数、さらには従事者や生産設備などについても統計原単位情報として報告徴収される場合もありうる。しかしそれが表式調査の形態で徴収される限り、それらはそれぞれ別途積み上げられるだけである。従って、表式調査の結果数値からは、各地域集計レベルでの平均的な生産性のようなものは算定できても、個体ベースの情報が存在しないことから、属性別のクロス集計により生産の構造的特性や分布特性は得ることができない。表式調査は、全国計あるいは地域別の単純集計値の獲得という結果データ面での制約を持つ。

このような制約を打破し、より豊富な統計情報を獲得する目的で導入されたのが個票調査である。

表式調査でも収集すべき統計原単位情報の性格について適切に企画されていれば、静態量と動態量の混在といった混乱は回避できる。その意味では、初期の「物産表」に見られたこういった問題は、表式調査に固有の問題ではない。それでは、表式調査と個票調査の調査方法上の質的な差異はどこにあるのだろうか。

## 2) 個票調査と集計制約

個票調査の生命は、統計原単位情報が個体ベースで把握されていることが可能にする多様な組み合わせ集計の可能性にある。その中には現実的に意味のない組み合わせも存在するが、原理的には、フェイス項目も含め、全ての調査事項がいわゆるクロス集計事項の対象となりうる。

しかしながら、実際の調査結果の提供は、そのうちの一部に限られてきた。その最大の制約条件は、技術的な集計能力であった。戦後しばらくの間は、政府が調査等で獲得した統計原単位情報の集計作業に動員できる大型集計機械の処理能力は現在とは比較にならないものであり、従って製表される調査項目も自ずと限定されたものとならざるをえなかった。「統計法」第7条が、指定統計調査の集計事項及び集計方法についての承認義務を調査実施機関に課しているのは、このような技術的制約と無関係ではないであろう。第2の制約は、結果数字の安定性という統計技術的なものである。多重クロスの場合、作り出されるセルの数は組み合わせ変数の次数に従っていわば幾何級数的に増加し、それだけ個々のセルに落ちるケースの数は減少する。調査結果から意味ある安定的な傾向を読みとるには、ある程度まとまったケース数が求められ、また結果表の表示形式が複雑となることから、自ずと組み合わせられるべき集計事項の数も限定的なものとならざるを得ない。しかしながら、実際の結果公表は、次の事情により、より低位の次数のものとなる。それは、結果表の公表形態による制約である。公表は長い間原則として印刷物の形で行われてきたことから、印刷コストが多様な集計結果の提供の制約となってきた。

より高性能な集計機械の導入により、集計能力と他の結果公表の制約条件との間のギャップを次第に大きくしてきた。特に公表形態による制約条件については、集計結果の一部がコンピュータによる打ち出し結果の閲覧という形で公表範囲が徐々に拡大されてきた。なお、CD-ROM等の磁気媒体で提供されるものには印刷報告書にはない結果表が一部収録されているケースもある。

収集された統計原単位情報が有する情報量と公表提供される集計結果とのギャップを埋めるもう一つの方策として採用されてきたのが、いわゆるオーダーメイド集計である。このオーダーメイド集計では、所定の手続きにより申請承認された集計計画について、統計作成機関内部で再集計作業が行われ、結果表が申請者に提供される。

## 3. 統計原単位情報の提供

### 1) 目的外使用制度

集計表という形式での調査結果の提供については、既存の統計制度の基本的枠組みの中で、提供の範囲を徐々に拡大する方向でこれまで対応がなされてきた。とはいえ、行政あるいは研究者等の利用目的は極めて多様で、個々の利用目的の全てに対応した結果の提供について統計作成者側で対応することは難しい。一般に公表される既存集計がカバーし得

ない利用に対処するためにわが国の「統計法」は、指定統計調査の調査票の目的外使用制度を規定している。

「統計法」は、統計調査によって収集された統計原単位情報について、「何人も、…統計上の目的以外に使用してはならない」（第15条）と規定している。この条項に記された「統計上」については「当該統計上」〔(3)p.65〕と解釈され、これまでその運用がなされてきた。すなわち、ここでの「(当該) 統計上の目的」は、統計作成機関が個々の統計調査施行細則により予め公示した結果表の作成に限定され、事前に公示されていない結果表の作成も「統計上の目的」からは排除されるということになる。

そこで、このような集計目的での利用も含め、既存の集計結果がカバーし得ない統計原単位情報のいわゆる第二次利用については、同じく第15条第2項が、「総務大臣の承認を得て使用の目的を公示したものについて」の適用除外を規定している。なお、承認統計調査、届出統計調査の調査票の目的外使用については、「統計法」第15条の2第2項で、「実施者が、被調査者又は報告を求められた者を識別できない方法で調査票又は統計報告を使用し、又は使用させることを妨げるものではない。」と規定され、指定統計調査に準じた取り扱いが調査実施者において行われている〔(20)p.59〕。

わが国における目的外使用の承認手続きは、現在「指定統計調査調査票の統計目的外使用の承認申請に関する事務処理要項」（昭和40年2月行政管理庁長官決定）に従って行われている。

個体ベースである統計原単位情報（磁気媒体上に記録された個票イメージでのデータ）の場合、例えば氏名や企業名、住所、電話番号などの個体識別情報を削除済みのものであっても、個体が特定されるリスクは集計量とは比較にならないほど大きい。他方で、「統計法」は、できるだけ正確な統計原単位情報の収集のために、申告者（被調査者）に調査における秘密の保護を約束している。統計原単位情報の使用目的について上記のような厳密な運用が行われてきたのは、調査に伴う秘密の漏洩が統計調査そのものの根幹に関わることから、調査協力者のプライバシーの保護を至上命題とするとの判断に基づくものといえる。

とはいえ、このような運用が収集される統計原単位情報の利活用という点では、大きな制約条件となっていることも事実である。なぜなら、統計作成機関そのものによる追加的な製表、あるいは地方自治体や他の政府機関による作表作業についても、目的外申請制度による許可申請が求められることになる。さらに政府以外の統計利用者が目的外制度による統計原単位情報の使用を希望する場合、その利用目的の公益性が求められる。この点で例えば学術的な研究については、高度に専門的でしかも公益性が高いことが承認の要件とされ、純粋に学術的な研究への利用は公益性を充足していないとして承認の対象から除外する形で制度の運用がなされてきた。

## 2) 目的外使用制度の運用実態

表1は、1985年以降の目的外申請の許可件数の申請主体別の推移を示したものである。

表1 指定統計調査票の使用主体別許可件数

	国	地方公共 団体	大学	その他
1980年	80	112	16	23
1981年	65	103	28	26
1982年	65	93	19	18
1983年	61	106	28	28
1984年	58	62	18	16
1985年	58	64	15	13
1986年	70	90	15	19
1987年	52	79	18	13
1988年	52	50	7	7
1989年	49	64	4	8
1990年	60	54	10	7
1991年	51	54	7	9
1992年	50	57	8	14
1993年	36	46	15	15
1994年	51	45	7	6
1995年	58	77	8	13
1996年	79	63	18	7
1997年	84	43	35	11
1998年	86	62	23	11
1999年	91	51	20	17
2000年	68	51	22	12
2001年	71	43	12	10
2002年	86	75	15	36

注)「その他」は、その他の公共団体、国公立研究医療

資料出所:『統計基準年報』、『統計情報』より作成

このデータから、目的外使用許可状況についていくつかの特徴を読みとることができる。

まず、許可数の大半が国と地方公共団体からの申請に対して出されており、ここに掲げた年次についていえば、その割合は実に8割を超える。これに対して大学からの申請許可は、全体の1割にとどまっている。特に大学から出された申請に対する許可件数は、特徴的な動きをしている。すなわち、1980年代半ば過ぎまでは毎年20件前後と比較的多数の申請が許可されている。しかし、如何なる事情によるものかは明らかにされていないが、その後90年代半ばにかけてはその数が激減している。さらに1997年以降は、その許可件数は再び大きく増加することになるが、この理由については後に触れることにしよう。

### 3) 目的外利用申請の審査とその問題点

実際の利用申請にあたっては、例えば再集計については、作成する個々の表の表頭、表側の個々の分類区分、回帰分析等の統計解析手法を使用する場合には変数として利用する統計項目の名称を明記した詳細な利用計画の提出が求められる。

目的外申請制度の運用を所管する総務省統計局統計基準部で受け付けられた申請書類は、利用目的が公益性の要件を充足しているかどうかについての点検が行なわれる。公益性要

件が満足されている申請案件について、申請のあった当該調査を所管する主管課にまわされ、利用計画の内容についての第一次審査が行われる。そこでは、個体が識別されるおそれがないか、利用計画での表記が調査資料のそれと適切に対応しているか、また得られる結果が十分に安定的かどうかといったような観点から審査され、不適切な部分については修正あるいは削除が求められる。調査主管課による審査を経た申請については、統計基準部が利用目的と再集計等の利用計画との整合性を審査し、利用者本人並びに所属長による調査資料の取り扱い・保管の責任体制についての書面での確約をとった上で最終的な使用承認を与える。

このような審査過程をとることから、審査受付から使用承認の『官報』告示により調査資料が利用可能となるまでに早くとも半年以上必要となる。審査に長時間必要とされることから、単年度主義で運用される各種の研究補助金による研究とは整合的でなく、利用者からより迅速な審査を求める要望がだされてきた。また、利用が承認された場合にも、利用期間が一般に半年程度と短く、また調査資料のデータ処理は許可された利用計画に従って行うものとされ、試行錯誤的な利用さらには集計変数の一部を除外することで多重クロス次元を減じた集計も認められないという点にも利用上の制約が大きく、行政機関あるいは地方自治体による利用申請に対して行われているような包括承認による調査資料の提供が求められている。

## II. ミクロデータ提供の現状

### 1. 諸外国におけるミクロデータの提供

わが国では個体ベースの統計資料の提供・利用は、基本的に目的外使用申請制度という形で今日に至っている。これに対して諸外国では、調査資料から再抽出された個体ベースの調査資料に、個体識別情報の削除、地域や職産分類についての区分の統合、年齢や所得区分での上位クラスの統合 (top coding)、といった一連の匿名化措置を施した個体ベースの中間フィル (ミクロデータ) を作成し、公共利用ファイル (public use files: PUF) として一般の利用に供するという形で、すでに早いところでは 1960 年代から利用が行われてきた。以下に、欧米を中心とする主要国におけるミクロデータ提供の現状を概観しておこう。

#### ①アメリカ合衆国

アメリカのセンサス法<sup>(註3)</sup>はその第9条で、「この法律の規定に基づいて提供された情報を統計目的以外に使用すること、そのデータを提供した特定の事業所又は個人が識別されるような形で公表すること、あるいは商務省、局あるいは機関によって認定された職員以外の人々が個々の報告を研究することを許可してはならない」と規定している。このよう



にセンサス法そのものにはマイクロデータの提供を明示的に認める条文は存在しないが、センサス局の活動を規定した合衆国法典 (US Code) 第 15 編が個別データの提供の法的根拠を与えているとされ [(23)p. 791]、実務上はセンサスの種類毎に調査結果の公表についての規定があり、それに従ってマイクロデータの作成、提供が行われている [(10)p. 5]。集計結果表だけでなく非集計データであるマイクロデータについても積極的に提供していくという同国の政府統計機関の姿勢には、自らをデータの提供者とユーザーの仲介者であり調停者として位置づけるという発想がその根底に横たわっている [(22)p. 91]。

同国におけるマイクロデータの提供は、センサス局がオーダーメイド集計の作業負担の軽減や研究者からの要望を受ける形で 1963 年に 1960 年人口住宅センサスの詳細調査票 (longer form) から作成した 1/1,000 抽出標本データを提供したのをもって嚆矢とする [(10)p. 5、(11)p. 2]。その後、1970 年センサスマイクロデータについては抽出率は 1% に引き上げられ、1990 年センサスでは 5% 抽出標本データが、また 65 歳以上の高齢者がいる世帯に関してもマイクロデータファイルが作成され、一般に提供されている。

センサスマイクロデータ以外では、経常人口調査 (Current Population Survey : CPS) が 1960 年代から作成、提供され、現在では所得・各種給付プログラム参加調査 (Survey of Income and Program Participation: SIPP) を初めとして、人口や世帯に関する殆どの統計調査について、センサス局はマイクロデータの提供を行っている。また、製造業調査 (毎年) および製造業経済センサス (5 年毎) についても、特に前者についてはパネル化したマイクロデータ (Longitudinal Research Panel: LRP) も作成されているが、事業所データについては、秘匿性の確保という点から、後述の特別宣誓職員に使用者を限定して提供されている [(11)p. 11]。

センサス局以外では、国立健康統計センター (National Center for Health Statistics: HCHS) が National Death Index を、また社会保険庁 (Social Security Administration: SSA) や国立教育統計センター (National Center for Education Statistics: NCES) もマイクロデータを作成、提供している。

アメリカでは、主に 2 つの方式でマイクロデータが提供されている。一つは、大学や研究機関等の研究者さらには民間企業などの一般利用者に提供されるもので、Public Use Microdata (PUM) と呼ばれ、その用途並びに使用方法には一切の限定は付されていない。なお、PUM については一般に市販されており、海外にも提供される。学術研究目的でのマイクロデータの提供は、大学間共同利用組織であるデータ・アーカイブ・センターを経由して行われることが多いが、特にミシガン大学の (The Inter-university Consortium for Political and Social Research: ICPSR) には、大学をはじめとする民間のマイクロデータだけでなく、経常人口調査 (CPS) や消費支出調査 (Consumer Expenditure Survey: CES) といった政府統計マイクロデータも登録され、広く利用されている。ミシガン大学社会研究所が連邦政府の資金的援助を受けて実施している所得動学パネル調査 (Panel Study of Income Dynamics: PSID) は、経済理論仮説の検証に多方面で利用されている [(16) p. 186]。

もう一つは、特別宣誓職員(a special sworn employee)としてセンサス局職員と同じ守秘義務を課することにより、企業・事業所データのようにより個体識別可能性の高いマイクロデータへのアクセスを認めるものである。なお、後者の場合には、マイクロデータの解析処理は、原則として当該データ保有組織内部に限定される(in house use)が、ボストンやピッツバーグなどにリサーチ・データ・センターを開設し、遠隔地の利用者の便宜をはかっている〔(10)p. 5〕。センサス局の他、歳入庁所得統計部、国立教育統計センターにはリサーチフェローの予算が計上されており、研究目的が庁局の業務目的に適合すると認められる研究に対してはフェローシップが提供され、宣誓職員としてこの種のマイクロデータを利用することができる仕組みになっている〔(11)p. 10〕。

## ②カナダ

アメリカにおけるマイクロデータの提供に触発され、1960年代初頭、カナダでもこの種のデータ提供への要望が多方面から寄せられた。それを受けてカナダ統計局では、当初、局内に特別組織を設置し、オーダーメイド型のマイクロデータに基づく解析処理サービスを開始した。しかし、サービス提供に時間がかかり、利用者がデータに対して対話形式で処理ができないことから弾力性に欠け、費用が高いといった一連の問題点が露見することになった〔(11)p. 17〕。

この問題に抜本的に対応するために、カナダ統計局は1971年に統計法を改訂し、マイクロデータ提供に対する法制度的な枠組みの整備を行った。すなわち、個票情報の取り扱いに関して「情報の提供者から事前の承諾を得ないで公表することを全面的に禁止」していた旧法を改正し、「統計局の職員や雇用者で、宣誓を行った者を除き、何人に対しても識別可能な個々の調査票の閲覧を許してはならず、宣誓を行った者は、何人もまた手段の如何を問わず、個々の調査票から得られた事項が特定の個人、個別企業又は個別の団体であると識別できるような方法によって公表又は故意に開示してはならない」(第17条)とした。これは、「情報の利用という点で不必要な制約を課」していた制度を改め、「カナダの法規を国際的な慣行」に追いつかせるものであった〔(11)pp. 18-19〕。

このような制度的枠組みの下にカナダでは、人口センサスについては、1%と2%抽出の標本マイクロデータが、また労働力調査、所得調査、総合社会調査、労働及び所得ダイナミックス調査についてマイクロデータが作成され、提供されている。

政府機関については、「マイクロデータ使用に関する了解覚え書き」を統計局との間で取り交わすことで利用でき、政府機関以外の利用者については、「マイクロデータ使用許可書」への署名が求められる。なおカナダでは緊縮財政の制約下で、統計の作成にも原価主義が導入され、その関係で提供されるマイクロデータもかなり高い価格設定がされた。そこで学術研究利用者側では大学その他の研究機関を「カナダマイクロデータ利用者協会」に組織化することで個々の利用者のマイクロデータ利用の便宜が図られている〔(11)p. 20〕。

## ③イギリス

イギリスでは、個票データの提供に関する明文化された規定はなく、その提供による秘

匿性の侵害行為がないというこれまでの運用実績を踏まえて、慣習法的にその提供が行われている。

イギリスにおける政府統計マイクロデータの提供は、統計データの種類並びに提供方式を異にする3つの部分から構成される。サーベイ・マイクロデータ、センサス・マイクロデータそしてLSデータがそれである。

エセックス大学に研究・教育を目的とする電子データの保管と利用のためのアーカイブが設置されたのは1967年である。アーカイブでは、開設当初、研究者などが作成した非政府統計を収集していたが、1980年代からは、政府が実施する一部の標本調査についてもマイクロデータの収集保管、提供活動を行ってきた。

アーカイブはまた1976年に社会科学研究および教育へのデータ利用を目的に設立された欧州社会科学データ・アーカイブ協議会(Council of European Social Science Data Archives: CESSDA<sup>(注4)</sup>)の英国における該当組織となっている。

一方、イギリスではセンサス・マイクロデータは約10年の歴史しか持っていない。1993年に国家統計局が1991年人口センサスから作成した2%個人匿名標本データ(Samples of Anonymised Record: SAR)と1%世帯SARがそれである。これらは、センサス全体の約1割にあたる完全コード個票(fully coded records)から再抽出により作成された。

エッセックスのアーカイブが保有するマイクロデータ場合、使用にあたっては利用者本人の署名のみが求められるのに対し、SARの使用にあたっては、所属機関長がEnd-user License Agreement (EULA)に署名した上で利用者本人がUser registration documentに署名しなければならない。なお、商業利用あるいは非学術的目的にSARの使用を希望する利用者は、Commercial End-user License Agreementへの署名が義務づけられている。

SARの提供対象は原則として連合王国(UK)域内の研究者、政府機関勤務者、商業利用者、学生とされているが、最近ではイギリスの高等教育機関の研究者との共同研究という形で外国の研究者の利用にも道が開かれるようになっている。

イギリスの政府統計マイクロデータで特に特異な存在であるのが、LSデータ(Longitudinal Study Data)である。これは、職業別の死因分析や女子の出産行動などの分析を主な目的に、1970年代初頭に旧人口センサス調査局(Office of Population Census and Surveys: OPCS)が1971年センサスの約1%標本抽出個票データをイングランド・ウエールズ登録局が保有する政府保健中央登録(The National Health Service Central Register)データとのリンケージにより作り上げられたマイクロデータである。LSデータは、年間に均等に配置された4誕生日(LS Birthdays)の該当者について、その後もセンサス実施の度にいわゆる縦断的データとして更新され、すでに30年間にわたるパネルデータとなっている。なお、LSサンプル該当者に対するデータリンクについての告知はされていない。

LSデータについては個人に関する秘密事項を数多く含み、データ上、個体ベースで追跡することから、個体識別される可能性がマイクロデータの中でも特に大きい。このため、局内の他部門からの申請も含め、政府の他の諸機関や研究者から出された利用申請につい

ては、使用目的の公益性、データの秘匿性が侵害されないかどうかについての審査が国家統計局に設けられた審査委員会（Research Board）で行われる。それに合格した利用申請については、局内のLS課との共同プロジェクトが立ち上げられる。なお、LSデータのハンドリングはLS課内の一部の職員に限られ、申請者には分析結果だけが提供される。

#### ④ドイツ

ドイツでは、1971年のマイクロセンサスのマイクロデータが学術研究に対して提供されたのを契機に、社会経済の構成要素である個人や世帯、階層に関する個別データを多変量解析、パネル分析、縦断面（ロンジチューディナル）分析が広範な広がりを持って展開され、それらは今日、マイクロ分析として社会科学において一つの研究領域として確固とした地位を獲得するに至っている〔(8)p.8〕。マイクロ分析は、主に、社会的不平等と貧困、教育の機会均等の社会階層移動への影響や女性の社会参加というような公共性の高い社会福祉面での政策課題の選択ならびに政策評価を対象としている。

1987年の連邦統計法の改正により、連邦統計局及び州統計部局は、個々のデータの識別が過度の時間、費用及び人員によってのみ可能であり、かつ受領者が在職者、公務のために特に宣誓した者又は第7条に定める義務を負う者である場合に限るとの限定つきながら、「科学プロジェクトの用に供するため、高等教育機関又は独立の科学研究任務を委任されているその他の機関に対し、個々のデータを伝送することができる」（第16条(6)）ことが制度化された。これに伴い、マンハイム大学 Walter Muller 教授の指揮の下、個体の識別に多大な時間、費用、労力を要する「事実上の匿名化」の方法に関する大規模な調査プロジェクトが1988～91年にかけて連邦統計局、州統計局、連邦及び州のデータ保護コミッションナー、マンハイム大学、マンハイム調査分析センター（ZUMA）によって実施された〔(19)pp.174-5〕。

ドイツにおけるマイクロデータの学術研究向けの提供窓口となっているZUMAでは、1962年以降の多くの年次のマイクロセンサス、1983年所得消費調査など多くのマイクロデータを保有しており、1989年からは連邦統計情報システム（STATIS-BUND）を通してマイクロセンサスデータを利用できるようになっている〔(7)p.43〕。また、1997年には、初めて連邦統計局のマイクロセンサスのマイクロデータが科学目的ファイルとして有償（130マルク）で一般に提供されるようになった〔(8)p.9〕。このファイルは、抽出率1%で実施される人口・社会に関して約160項目を持つ年次ベースの標本調査の個票データから70～80%の再抽出リサンプリング・データとして作成されるものである。この他にもドイツでは、5年周期で調査が実施される所得・消費調査（抽出率0.3%の世帯調査）について1998年に科学目的のマイクロデータファイルが、また雇用統計についても1%の再抽出率で、さらに、1991、92年病院統計、1995年建物・住宅センサス、1991年以降の道路交通事故統計の事実上の匿名化マイクロデータの提供が行われている。

#### ⑤オーストラリア

オーストラリアでは1970年代までは、印刷物による集計結果表の公表を原則とし、追加

的な集計要望については、いわゆるオーダーメイド方式による特別集計の提供という形での対応がなされてきた。利用者によるニーズの高まり、多様化に対応しきれなくなったことから、オーストラリア統計局は、1983年に統計法(Census and Statistics Act)を改訂することで、マイクロデータが提供できる制度的枠組みの整備を図った。

改正前の1973年法では、報告者の秘密保護に関して「統計局長、局職員は…本法に基づいて提供された個票のいかなる内容、情報も漏らしてはならない」(第24条)と規定されていたものが、「統計局長又は統計局職員である者又は職員であった者は、規則(Determination)又は本法による場合を除き、提供されたいかなる情報も直接又は間接にいかなる者に対しても漏洩又は伝達してはならない」(第19条)と改訂され、規則第7条において、氏名、住所等の識別情報の削除、使用者の文書による誓約を条件に統計局長が個別統計データ(Individual statistical data)を開示できることが規定されている〔(12)pp. 50-51〕。

このような制度改革を受けて同国では、1980年代半ばに国民健康調査(National Health Survey)、所得住宅調査(Income and Housing Survey)のマイクロデータが提供された。またセンサス・マイクロデータについても、1981年調査に基づくマイクロデータが、ほぼ同時期に提供されている〔(12)pp. 49-50〕。

マイクロデータの利用申請は、統計局内の個々の部局長あるいは次長を通じて所定の書式に従って行われ、個々の利用申請に関わるマイクロデータ提供の可否の判定は、統計局内に設置されたマイクロデータ検討委員会に委ねられている。提供が許可された場合、有料で利用者に提供される。なお、企業・事業所関係のマイクロデータは提供されていない。

#### ⑥ニュージーランド

ニュージーランドでは、1975年の統計法改正によって、地方政府を含め、政府機関が共同実施した統計調査について個票を関係機関が相互に使用できること、ならびに他の政府機関がその機関の任務遂行に必要な研究ならびに統計目的のためのマイクロデータの提供が認められた。

一方、政府機関以外の利用者からの利用申請に対しては明確な法的規定がなく、統計法の例外的な運用としてケースバイケースで認められてきただけである。1996年に政府機関以外からの利用申請への対応を定めた手続き規定(Protocols-Micro Data Access Management)が設けられたが、統計法上明文化された規定がないままその拡張解釈として今日に至っている。

#### ⑦オランダ

オランダでは、「中央統計局及び統計委員会の設立に関する法律」の第11条が、同局が職務の遂行に関連して収集したデータは統計目的のためだけに使用することができる旨を規定しているが、同じく第13条によって、個体識別されないための秘密保護措置がとられている場合に限り、大学その他の研究機関、中央計画庁、EUROSTAT及び中央統計委員会が承認した組織等に提供できることを規定している〔(9)p. 11〕。

オランダにおけるマイクロデータの提供は1970年代に開始された。1985年までは中央統計局からデータの提供を受け国立データ・アーカイブ(National Data Archive)が提供を行ってきたが、秘匿性の確保の問題からそれ以降は中央統計局が直接提供することになった〔(9)p. 11〕。データ提供に当たっての価格の問題に対処するためにオランダ科学研究機構(The Netherlands Organization of Scientific Research: NOW)が仲介機関<sup>(注5)</sup>となってマイクロデータ提供の便宜を図っている〔(9)p. 11〕。

労働力調査、社会生活調査、家計調査については公開利用ファイル(Public Use Files)が提供されているが、その他にも契約ベースでの使用システムもある。後者の場合には、地域変数が詳細であるなど、PUFに比べてデータの内容がより豊富である。なお、契約ベースでのマイクロデータの利用に当たっては、利用希望者本人が秘密保護に関する誓約書にサインするとともに、利用者が所属する機関と中央統計局との間で利用に関する契約の締結が求められる〔(9)p. 11〕。

#### ⑧フランス

フランスでは、世帯関係のマイクロデータとして、人口センサス、労働力調査、生活調査、家計調査について、公共利用データ(Public Use Data)が提供されている。なお、企業事業所データについては秘匿性確保のため、利用に当たっては国家情報統計委員会(CNIS)の承認による利用契約の締結が求められる〔(22)p. 787〕。

#### ⑨韓国

韓国では、統計法(1962年1月法律第980号、1999年1月最終改正)第13条第2項が秘密保護規定の一環として、「統計作成のために収集された個人、法人又は団体に属する秘匿を要する基本資料は、これを統計作成以外のいかなる目的にも使用してはならない」と規定している。しかし同時に他方で、<統計データの普及>を定めた第16条では、統計機関の長は、本法第13条の規定に反しない限りにおいて、大統領令が定める条件に基づいて統計データを広範に使用させるものとする」に依拠して、近年、マイクロデータの提供も行われている。

韓国統計庁が実施したセンサスや標本調査の個別情報は、個人や事業所の秘密が厳格に保護される限りでそれを必要とする利用者に提供され、利用者の求めに応じたオーダーメイド集計サービスも提供されている。特にマイクロデータの提供と関連して注目される点は、2000年5月より、人口・住宅センサスの2%ファイルがいわゆる公共利用ファイルとして一般にCD-ROM提供されるようになった点である〔(18) p.471〕。この他にも韓国統計庁では、利用者が同庁が作成した個別データに庁内でアクセス、処理ができる”on-site access” systemもすでに稼働している〔(18) p.471〕。

このように、マイクロデータの提供は国によって、またデータの種類によっても異なる。表2は、主要国のマイクロデータの提供方式を示したものである。

表2 各国におけるマイクロデータの提供方式

提 供 方 式	採用している国
<p><b>公開用データ(秘匿処理済データ)の作成、提供</b></p> <ul style="list-style-type: none"> <li>・統計作成部局等で秘匿処理を施した公開用データを作成し、利用者はそれを用いて集計</li> <li>・自由に使用できる国と誓約書が必要な国あり</li> <li>・企業関連統計はこの方式では提供できない</li> </ul>	アメリカ、カナダ、イギリス、フランス、ドイツ、オランダ、オーストラリア、ニュージーランド、韓国
<p><b>インサイト利用(宣誓職員方式)</b></p> <ul style="list-style-type: none"> <li>・利用者は秘密を守ることを宣誓し、統計局等のサイト内でのみ利用可(法的な秘密保持義務を負う)</li> <li>・統計局との共同研究のみに限定していることあり</li> </ul>	アメリカ、デンマーク、ニュージーランド、スウェーデン、イギリス (LS)、韓国
<p><b>契約方式(目的外使用申請方式)</b></p> <ul style="list-style-type: none"> <li>・マイクロデータをそのまま利用者に渡すが、原データを直接利用させるのではなく、何らかの秘匿措置を講じることが多い</li> <li>・利用者を政府機関等に限定していることあり</li> </ul>	日本、アメリカ、イギリス、フランス、ドイツ、ノルウェー、オランダ
<p><b>オーダーメイド集計方式</b></p> <ul style="list-style-type: none"> <li>・統計作成部局等が、利用者から提出された集計様式でプログラムを作成して集計又は利用者が作成したプログラムで集計し、結果のみを提供(原則、有料)</li> </ul>	カナダ、オランダ、イギリス、オーストラリア、韓国

## 2. わが国におけるマイクロデータの提供

わが国では、1985年以來統計審議会がほぼ10年ごとに統計行政の中長期的指針として政府統計の在り方についての答申を提言してきた。答申でマイクロデータにはじめて言及されたのは、1995年の『統計行政の新中・長期構想』が標本データ（個票データから必要に応じて抽出を行い、地域区分や世帯番号等の個体の識別子を消去するなど個体の識別を不可能にしたもの）の提供についての課題設定を行ったのが最初である。そこでは、「統計調査の結果」が「多額の費用と国民の協力を得て収集された貴重な情報資源」であり「国民共有の財産として社会全体で活用すべき」との立場から、「分析能力を高めたユーザーが、これまでのように結果表中心の統計利用から小地域、小集団の統計にまで感心を深めている状況を踏まえ、統計作成部局は秘密の保護に十分な措置を講ずるなど国民のプライバシーに配慮した上で、指定統計調査票の統計目的以外への使用（以下「目的外使用」という。）の積極的な活用や標本データ等の提供の検討により、ユーザーの多様な統計ニーズにこたえていく必要がある。」〔(5)p. 60〕と指摘されている。

また、同答申で標本データの提供については、「個体の秘密保護の担保方策を中心に、外国の制度及び提供例、国内外におけるニーズの実態、現行法制度との関係、具体的な提供方策等について、おおむね2～3年を目途に専門的、技術的な研究を行う必要がある」〔(5)p. 75〕とされており、これを受けて行政では答申のフォローアップ委員会が設置され

専門的、技術的な研究が行われた。その結果、一部の標本データについては、その実用性が見込まれることが検証された。

他方、マイクロデータの潜在的な最大の利用者である研究者の側でも、1996年に3年間継続の特定領域研究プロジェクトとして、「統計情報活用のフロンティアの拡大ーマイクロデータによる社会構造解析」（領域代表者：松田芳郎現東京国際大学教授）が立ち上げられ、秘匿性確保の統計技法の研究、海外におけるマイクロデータ提供の法制度の研究、そして個票ベースのデータによる実証という3つの研究領域にわたっての精力的調査研究が行われた。表2で1997年から2000年にかけて大学から出された目的外利用申請に対する許可件数が急増しているのは、同プロジェクトの実施に伴う申請許可によるものである。さらに、この特定領域研究の成果を踏まえ、2000年から2004年の5カ年計画で（財）統計情報研究開発センターを事務局として「マイクロ統計データ活用研究会」が組織され、ここを窓口としてリサンプリング・データのユーザビリティの研究を目的にいくつかの政府標本調査からの再抽出標本から作成されたリサンプリング・データの研究面への提供が行われてきた。

ところで、「審議会等の整理合理化に関する基本的計画」（平成11年4月閣議決定）を受けて統計審議会は2001年1月より法令により審議会等への必要的付議が定められている事項についてのみを審議事項とする「法施行型審議会」に移行した。その結果、統計審議会は、それまで保有していた政策審議・基準作成機能を喪失することになった。このため、統計行政の新たな中長期的な指針は、各府省統計主管部局長等会議において取り纏められることになった。

マイクロデータの提供が個体ベースのデータのより広範囲な提供という既存の制度運用の枠組みを超えるという側面を含むことから、1995年の答申を受けた取り組みは慎重に進められた。そして2003年6月に纏められた新たな答申「統計行政の新たな展開方向」では、匿名標本データの作成・利用について次のような形で取り纏められている。

まずその作成・利用については、「秘密の保護を前提に、匿名標本データに対する社会的コンセンサスの必要性、法令上の諸問題等に関するこれまでの検討結果を吟味した上で、実用化に向けて検討する」〔(20)p. 58〕という慎重な表現となっている。

またそこでは、匿名標本データという形でのマイクロデータの提供の実現に向けての具体的な方策として、次の3点が指摘されている。

- ①関係府省間で、これまでの検討結果を踏まえ、匿名標本データの定義及び匿名標本データの作成になじむ統計調査の範囲を明確にするとともに、平成15年度中に制度面を含め匿名標本データの作成・利用に当たっての課題を整理する。その際、匿名標本データの導入する場合の利用方法、利用場所等についても検討する。
- ②関係府省は、この整理を踏まえ、平成16年度以降、順次、匿名標本データの作成・利用のためのガイドラインの作成など実用化に当たっての基本的枠組みを検討・構築する。
- ③関係府省は、当面、匿名標本データの利用ニーズや導入に当たっての課題等を把握



するため、必要に応じ試験的な匿名標本データの作成を行う〔(20)pp. 58-59〕。

このように、わが国におけるマイクロデータの提供・利用システムの構築に向けての行政側の取り組みは、多くの国がマイクロデータ提供の制度整備のために統計法規の改正などに積極的に取り組み、そのような中でのマイクロデータの利用実績が新たな統計利用の地平を開拓している現状と比較すれば、なおかなりの温度差が感じられる。

### Ⅲ. マイクロデータが提供しうる追加情報の可能性

マイクロデータについては、秘匿性の保護の観点から各種の匿名化措置がデータに対して施されている。また原データセットから一般に再抽出して作成されたデータセットであるため、それが有し利用可能な情報量は、例えば包括承認によって使用が認められたデータセットの利用可能性に比較すればやや小さい。しかしそれへの一般の利用者からのアクセス可能性の広がりの中で、行政機関という特定の利用者だけを想定した包括承認とは質的に異なる。

このようにマイクロデータは、統計調査によって収集された統計原単位情報の一部を匿名化措置によって失ってはいるものの、少なくともその使用については包括承認による統計資料の提供と同等の使用可能性を持っている。そこで次に、集計量として提供されてきたデータに対するマイクロデータが与えうる新たな利用可能性についてここで簡単に整理しておこう。

#### 1. マイクロデータそれ自体による利用可能性の拡大

##### 1) 独自集計としての利用可能性

マイクロデータの場合、分析者が有する固有の分析目的に従って、集計過程そのものを自ら企画・実行することができる点が、既存の集計結果を前提とした統計利用と質的に異なる。

かつてと異なり、その後の情報処理技術の急速な進歩は、集計作業そのものについての時間、費用面での制約条件を本質的に緩和することになった。またデータの提供面でも、磁気媒体やインターネットの活用によりそのコストは印刷ベースと比較すれば格段に小さい。このような技術的条件の変化を受けて、公表される既存集計の量はかつてに比べれば大きく拡大している。他方で利用者の分析目的は極めて多様で、オーダーメイド集計を加えても既存の集計結果表がそれらの全てをカバーすることは現実的に不可能である。

まず、マイクロデータに固有な集計利用面でのメリットの第1としては、利用者が各自の問題感心に従って既存の集計結果にはない調査事項（変数）同士を組み合わせたクロス集計、あるいは既存の集計組み合わせにさらに変数を追加することで、既存集計からは得ら

れない結果表を手にできる点がある。このことは、既存集計が全く取り上げていない変数相互の関連の分析、あるいは既存集計になお複数の変数の作用が未分離の形で混在している場合に、他の要素の作用をコントロールして特定の要素の作用のみを純粹に析出する場合に有効となる。

マイクロデータの集計利用の第2のメリットは、分類区分の自由度の拡大である。既存の集計を利用する場合、結果表の分類基準が研究者の分析目的と必ずしも整合的でないケースが少なくない。既存の結果表を前提した従来の統計利用の場合、結果表に存在する分類区分をどのように統合すればより現実の特徴を反映した集計表が得られるかという「統計的グループ分け論」の形で主として立論が展開されてきた。既存の結果表を前提する限り、研究者は自らの分析目的に従った独自の結果表を既存の統計に対置することはできない。その結果、いわば理念的に定式化された集計方法に照らして、しかるべき結果表を政府統計が提示していない点を「政府統計の限界性」として統計批判が展開されてきた。

調査事項に記載された回答選択肢の分類区分がこの種の「提案」にも対応する形で調査票が設計され、統計原単位情報が収集されていれば、この種の統計批判は政府統計の限界性に必ずしも該当するものではない。従ってこのような形での統計批判は、集計過程のありよう次第では解消しうる性質の問題である。

さらにこのことと関連して、調査票に設けられた調査項目さらにはその回答選択肢に分析者の利用目的に正確に対応する概念設定が存在しない場合にも、マイクロデータの場合には、複数の調査項目あるいは回答選択肢を組み合わせることで、それに近似させた概念を独自に構成できる可能性も存在する。このような形で統計原単位情報の情報量の拡大をはかる方法は海外ではすでに実用化され、新たに追加された変数を付加した個体ベースでのデータが利用に供されている。例えば、英国国家統計局（ONS）が作成し、現在マンチェスター大学CCSR（Centre for Census and Survey Research）が提供の窓口となっているセンサス・マイクロデータでは、人口センサスの統計原単位情報のうちマイクロデータの変数に採用されている調査項目の他にも様々な変数が導出変数（derived variables）として個人及び世帯の個体レコードにそれぞれ追加され、提供されている。

## 2) 統計的手法の直接適用によるマイクロデータの利用可能性

マイクロデータの利用を集計結果の利用から質的に区別するのが、回帰分析等の統計解析手法の適用である。

多くの変数を持つ膨大な多重クロス表を重回帰分析によって代置した場合、関数式の定式化という要素が不問にできるという点ではクロス表の方に利があるものの、結果の読み取り、その意味付けという点では後者の利用可能性は大きい。すなわち、推定されたパラメータの検定は、従属変数に対する個々の独立変数の関与の有無についての判定基準を与え、さらにその大きさは従属変数に対する影響度の大きさを評価する基準を与える。また同一調査の横断面データ分析の比較からは、その間の構造変化の有無といった面にもアプローチすることができる。

回帰分析以外にも様々な多変量解析手法をマイクロデータに直接適用することによって、いろいろな形で調査項目間の相互関連を分析することができる。

## 2. データ・リンケージによる追加的情報の創造

調査個票に記載された統計原単位情報が他の情報と接合されるのにはいくつかのケースが考えられる。統計調査における把握精度向上のための対象者名簿との照合、あるいは企業・事業所関連統計の同一調査資料における企業単位での名寄せ集計のための個票データの接合も、広い意味ではデータ・リンケージにあたるといえる。しかし以下では、主として統計原単位情報の情報量を拡大する目的で行われる個票ベースでのデータの接合に限定してその特徴を論じることとする。

### 1) 完全マッチングと統計的マッチング

二種類の個票ベースでの統計原単位情報が接合される場合、接合における中心的ファイルは基準ファイル、また基準ファイルに接合されるファイルは参照ファイルと一般に呼ばれる。これらのファイルは、結合子あるいはマッチング変数といわれる共通情報、例えば世帯番号、事業所・企業番号、住所、氏名、年齢、資本金等を用いて接合される。

データ・リンケージによる調査間の接合には、同一の調査対象に関わる調査個票データを1対1で対応付け、相互に接合する完全マッチング（完全照合）と相互に統計的に同一と見なしうる属性を持っている者を相互に接合する統計的マッチング（統計的照合）とがある。

複数の調査で同一の調査対象に共通に設定した世帯番号を用いて個票データの接合を行った場合にも、例えば同一調査時点での年収など共通項目に乖離が発生するケースもある。これは調査期間の間に移転などして対象標本が差し替えられたり、一方の個票データが記載不備その他の理由で不的確として個票ファイルから削除された結果、照合すべき対が存在しないケースもあり得る〔(16)p. 244〕。

統計的マッチングは、結合する基準ファイルと参照ファイルに共通するマッチング変数値を用いて一種の距離を定義し、距離の小さいレコード同士を統計的に最も近いものとして相互に連結することに特徴がある。ファイルを相互に統計的マッチングによって連結するためには、マッチング変数の形式が予め統一されておく必要がある。すなわち、量的変数については数値そのものの調査するのが望ましいし、階級幅を持つ選択回答肢についてはその区分を共通に設定しておく必要がある。また質的変数の場合には、統一的な分類に従うコードの設定が必要となる。

統計的マッチングの特徴は、連結されたレコード同士が完全に同一の調査客体から提供された統計原単位情報であるという保証がない点にある。統計的マッチングについては、単に完全マッチングができない場合の次善の接合方法という消極的側面だけでなく、この統計的マッチングに固有の曖昧性が逆に、個体情報の秘匿という点で積極的な意義を持つ。

なぜならそれは、「マイクロデータを利用しても、調査回答者個別の名前には辿り着かないような開示形態」〔(15)p. 198〕として、マイクロデータの匿名性を補強し、調査結果の二次利用への道を開くという側面を持つからである。

## 2) 異種の統計間のデータ・リンケージ

『統計行政の新中・長期構想』は、調査事項の重複回避など主として調査客体の調査負担の軽減という視点からデータ・リンケージの推進が必要であるとして、企業・事業所関連統計調査について、「リンケージ項目（事業所の産業格付け、経営組織の区分、従業者内訳区分、資本金の階級区分等）の可能な限りの統一化、共通化を測ることとし、…中期的に検討を進め逐次実行に移す」〔(5)p. 14〕ことを提言している。

それとともに同答申は、新たな統計ニーズにこたえるための追加的な統計情報の獲得手段としてデータ・リンケージを位置づけ、「既存の統計調査の個票を利用するデータ・リンケージ手法や複数調査間のデータ・リンケージを前提」〔(5)p. 94〕にした統計調査の企画・設計の推進の必要性を指摘している。

わが国ですでに実施に移されている完全マッチングによるデータ・リンケージの例としては、かつての「家計調査」と「貯蓄動向調査」、「国民生活基礎調査」と同調査の調査対象世帯から抽出される諸統計調査との、また「工業統計調査」と「商鉱工業石油等消費構造統計調査」、「経済産業省企業活動基本調査」と総務省「科学技術調査」などがある。また米国では、統計調査の世帯情報に税務データをマッチングさせたケースなどがある〔(16)p. 239〕。

相前後して実施された調査個票データ同士の完全マッチングによるデータ・リンケージは、統計原単位情報の収集という意味では、結果的には両者を統合した調査事項を備えた調査票に従って実施された調査と同じ効果を持つ。これによって、新たな統計調査を企画・実施することなく、従って調査客体に追加的な調査負担を強いることなく、個体ベースのデータとして、解析可能な変数（統計調査項目）の拡大を実現することができる。それは新たな集計可能性を飛躍的に拡大することになる。

一方で予算制約と追加的な調査負担に対する調査客体側の忌避反応、他方で拡大する統計ニーズへの対応という相矛盾した要請を同時に実現できる手段として、データ・リンケージの持つ意味は大きい。

## 3) 縦断的データ・リンケージ

わが国では従来から個票データを過去に実施された同一調査の情報と連結して使用するという方法は、調査漏れや調査の記入内容のチェックする方法の一つとして用いられてきた〔(16)p. 206〕。しかしここでは、このような調査の精度改善のための調査技術としてのリンケージではなく、統計原単位情報の情報量を増やすことを目的としたリンケージに限定してその特徴を論ずることにする。

同一個体ベースで継続的に接続可能な情報としては、有価証券報告書がよく知られている。しかしこれは上場企業だけが対象でその範囲は限定されている。

わが国の政府統計で縦断的にデータをリンクさせロンジチュディナル・データを編成した例としては、「工業統計調査」を用いたものがいくつかある。まず統計審議会の調査技術開発部会は、エラーチェック・システム開発のために試験的に1981～85年の「工業統計調査」のある特定の県の個票データを用いてロンジチュディナル・データを作成している〔(15)p. 236〕。また、通産省調査統計部は産業空洞化の統計的把握に対する対応策として、外部委託調査研究として1985～95年の全数調査実施年について、全都道府県データを使用した、またこの他にも「通産省企業活動基本調査」からロンジチュディナル・データセットが編成された事例がある。しかしこれらはいずれも主として政府統計機関側での特定の利用目的との関連で作成されたものである。この点で目的を特定しない汎用型のデータベースとして作成されている諸外国におけるロンジチュディナルデータとは質的に異なる。

海外では、センサスや人口動態統計資料さらには各種の登録データをリンクすることでいくつかの国でロンジチュディナル・データが作成されている。すでに見たように、イギリスでは全人口の約100分の1の抽出率でNHSCR(National Health Service Central Register)の人口動態データと人口センサスの個票データとをリンクさせたLS(Longitudinal Study)データが1971年以降のセンサスデータについて作成されており、フランスでも1982年センサスを契機に1966年まで遡及させてLSデータと類似したEchantillon Demographique Permanente(EDP)が作成されている。またオランダでは、27,000人を対象にこの種のリンク統計が作成されており、アメリカやイタリアでも人口センサス等の静態データと動態データのリンクは、1970年代から試みられている。さらに国の統計作成の中心がレジスターベースで行われているデンマークやフィンランドでは、人口登録情報と人口動態に関する登録情報が全人口を対象に相互にリンクされている。

このような異時点で実施された既存の調査個票データの事後的リンクによるロンジチュディナル・データの編成とは別に、縦断的に同一の調査対象を継続的に追跡するものとしてパネル調査がある。

世帯を対象とした国際的に著名なパネル調査としては、アメリカではミシガン大学調査研究所が約4800世帯を対象に1968年以来継続的に調査されているPSID(The Panel Study of Income Dynamics)やオハイオ州立大学人的資源研究所・シカゴ大学のNLS(The National Longitudinal Surveys of Labor Market Experience)がある〔(16)p. 221〕。なお、わが国では財団法人家計経済研究所が、調査開始時点で24～34歳であった女性1,500人を対象に「消費生活に関するパネル調査」を1993年度から実施している。

現実のうちパネル調査というデータ形式によって初めて切り取ることが可能となる情報について、わが国の政府統計が関心を向けるようになったのは比較的最近である。現在のところ、わが国の政府統計でパネル調査として実施されているものとしては、厚生労働省統計情報部が2001年1月10日～17日、7月10日～17日に出生した新生児を対象に子供の成長・発達の様子や、子育てに関する環境や意識、行動の変化を把握する目的で開始した「21世紀出世児縦断調査」(承認統計2002年8月12日承認No. 23097～23100)があるだけ

である。

パネル・データを含む縦断的データは、多時点の横断面（クロスセクション）データと比較して次のような特徴を持つ。すなわち横断面データの場合、動的に変化する現実を一定時点でいわばスナップショット的に切り取ることで統計データが作成される。横断面データにおいても、例えば過去1年間の就業異動あるいは居住地や住居形態の変化などに関する調査項目を設けておけば、部分的にはあるが一定期間における位相の変化という動学的要素を統計として把握することができる。

この点で縦断面データの場合、多時点データが個体ベースで利用できることから、位相の変化パターンに従って多様な類型化が可能であり、それぞれの類型間の特性の異同についての分析ができる。例えば、 $t_0$ 時点で失業状態にあった者で $t_1$ 時点でも依然として失業状態にある者、 $t_1$ 時点では就業者あるいは非労働力人口として記録される者との間の就業動態に関して多様な類型化が可能であり、それら各類型毎に相互の属性に関する特性を比較することで、位相の違いをもたらしたと考えられる要因の検出を行うことができる。このように、分析にあたって動的視角からの集団構成によりそれらの特性分析を行うといった形での統計利用は、これまでのような多時点の孤立したクロスセクション・データでは果たし得ないものである。

同様の所得水準でしかも家族構成、年齢構成その他の世帯属性がほぼ等しい家計であっても、消費や貯蓄性向に大きな差異が見られるケースは少なくない。また企業でも、投入される労働量と資本量がほぼ等しいにもかかわらず、生産性に大きな差が生まれるケースがある。これらはいずれも統計調査項目として観測される要因ではうまく説明できない個体間の異質性に起因するものである。こういった要因の作用は、一般に「個体効果」あるいは「経営資源」と呼ばれている〔(16)p. 219〕。

既存の調査項目（変数）によって観測されないこのような個体効果については、クロスセクション・データによる分析ではいわゆる誤差項に含まれて処理されている。個体効果が他の変数と相関を持つ場合、誤差項に関する仮定が充足されないといういわゆる内性性の問題が発生する。回帰分析でクロスセクション・データと時系列データとで結果に違いが見られる原因の一つとして、このような個体効果が作用していると考えられる。

同一の個体についての多時点データとして得られるパネル・データの場合、少なくともその間に変化しなかった要素については個体効果の存在は捨象することができる。しかし実際の分析で文字通り個体ベースでの個体効果をモデル化すれば、いたずらにパラメータを増やす結果となり必ずしも生産的でない。このため、業種や家族類型といった集団別の「個体」効果をモデル化することで、説明変数の説明力についての評価が行われる。パネルデータにより個体効果を本来的な誤差項から分離することで、クロスセクション・データでは明らかにしえなかった分析結果がしばしば得られる。

## IV. 「公共財」としての政府統計データとその第二次的利用

### 1. 「公共財」としての政府統計データの提供

わが国における政府統計提供の現状を見るに、CD-ROMやインターネットといった新たな情報媒体や経路でのデータ提供が拡充されてきてはいるが、一般に公表されるデータの形式そのものは従来と同様、集計表のままである。

データの提供が集計結果表を原則としてきたのには、次のようないくつかの事情がある。

その一は、統計原単位情報の提供者である調査客体に対する秘密保護の担保である。「統計法」にも明記されているように、調査実施者である政府は、調査客体に対して統計原単位情報の提供を申告義務として強制する補償措置として、統計調査協力に伴う秘密の保護を保証している。集計量としての結果提供と比較すれば、個体ベースでのデータ提供の場合、個体（特定の個人、世帯、企業等）が識別される可能性は大きい。このことが、集計表を原則として、これまでの統計の結果提供が行われてきたものと考えられる。

第二の理由として、提供されるデータの安定性という側面がある。個々の統計原単位情報にはいろいろな形で調査誤差、標本誤差が含まれ、平均あるいは規則性から多かれ少なかれ乖離する。このため関係する他の全ての要因をコントロールしたとしても、当該変数（統計集計項目）について決して一点分布しているわけではない。そのような中で、集計し誤差を相殺することではじめて浮かび上がる規則性あるいは統計的法則をより安定的な形で表現することに表形式での調査結果表示の積極的な意義がある。統計的規則性はあくまでも集団現象の中に貫徹するもので、安定的な規則性は多くの個体からなる集計値からはじめて得られるという大数の原理が統計表による結果提供の背景となっている。

第三に、集計表による表現形式の簡潔性である。調査結果が表形式で提供される場合、何らかの統計解析結果としての提供あるいは個体ベースでのデータ提供と異なり、結果の読みとりには特別な専門知識を必ずしも必要としない。調査結果の公表が申告義務による調査強制の補償措置としての調査客体への結果の還元という意味合いを持つことから、いわば一般の利用者による汎用に最も適合的な調査結果の提供形態が集計結果表という形式であると考えられる。

政府統計機関にとって統計調査における「統計目的」とは集計結果表の獲得に他ならず、事前に公示された結果表作成以外の統計原単位情報の使用は、例え追加的な集計表の作成であっても本来の「統計目的外」の行為にあたる。統計原単位情報の使用方法についてのこのような政府統計機関側の認識、すなわち統計利用が集計量の利用を原則とするという認識がこれまで集計表という形での公表形態の根拠となってきた。このため、特に政府機関以外からの申請による目的外使用について、個票ベースでのデータが提供される場合にも、集計処理の結果として得られる集計表の個々のセルに落ちる数値が調査客体を識別するに至らないしかも安定的な比率ないし傾向を析出するのに十分なサイズのものであるか

否かについての実質的な審査が行われる。このような二重の視点からの審査の仕組みとして、目的外使用制度はこれまで運用されてきた。

近代調査統計の開始以来、調査実施者にとって集計表の作成が統計調査の最終到達点であった。このため調査実施機関では、記入済みの個票として調査客体から得られた統計原単位情報に対して、記入内容に関する論理チェック、集計に必要な符号（コード）化等の処理を施した上で、フェイス項目あるいは基本的な調査事項とのクロス表を作成し、結果報告書として公表してきた。その際、集計結果表に付記して調査結果の分析が掲載される場合にも、結果の概要あるいは調査の解説という形で調査結果についての若干の結果数字が紹介されるのが通例で、トピックス別にモノグラフシリーズ等の形で結果分析が公刊されるのも一部の調査にとどまっている。このように調査実施機関は、基本的に集計結果の提供者というスタンスを堅持している。その意味では、統計調査を通じて収集された統計原単位情報の第一次的使用は、調査実施機関による集計結果表の作表並びにその公表に他ならない。

## 2. 政府統計データの二次的利用

政府統計は、歴史的にもその大半が行政施策策定のための基礎資料を得ることを目的に作成されてきた。しかしそれが客観的事実を把握すべく実施されていることから、調査結果は、「行政施策策定の基礎資料」としての利用の枠を超えて、学術的研究や民間利用など多様な方面に利活用されてきた。このような、当該統計の作成機関以外の利用者による利用は、一般に政府統計の二次的利用(secondary analysis)と呼ばれている。

政府統計の二次的利用の場合、利用目的や分析上の関心事項が統計作成者のそれとは異なるのが一般的である。このため、統計作成者の調査目的から演繹的に導かれた統計の定義や分類基準さらにはそれに基づく集計概念は、二次的利用者の利用・分析目的と整合的でないことが多い。このような場合、統計作成者自身による第一次利用と異なり、政府統計の二次的利用者は、提供されている調査結果がその者の利用目的に適合的であるかどうかの吟味から出発しなければならない。かつて蜷川が『統計利用の基本問題』で提起した統計の批判的利用において、信頼性、正確性を機軸概念として既存の統計の作成過程の吟味の視座を提起したのは、まさにこのような観点からであった。その意味では、その延長線上に社会科学方法論として提起され展開されてきた統計学は、まさに統計の二次利用のための方法的規定であったといえる。

社会科学方法論説では、しかるべき社会科学に裏付けられた統計が本来反映すべき集団現象としての「大量」から演繹的に導き出される「大量の四要素」、さらにはそれを統計調査実施のために具体化した操作概念としての「大量観察の四要素」を軸に、一種の理念型として統計の作成過程が展開される。このようにして、本来得られるべき統計と現実に提供されている統計数字との理論的・概念的乖離については統計の「信頼性」、また調査・集



計過程にかかわる技術的な乖離については「正確性」として、既存の政府統計の吟味、批判が展開されてきた。

しかしながら、政府統計の対象となっているような全国ベースの調査による実際の統計作成には膨大な経費と人手が必要となる。このような統計作成の現実的条件を持ち合わせていないことから、社会科学方法論説は主体的な統計作成の指導理論としては、一部の研究者、労働組合あるいは一部の自治体など限定的な形でしか統計作成にまで具体化されず、多くの場合、対置すべき結果数字を持たない代案提案に留まらざるをえなかった。

「ありうべき数字」が現実的形態をとりえないことから、実際に提供されている政府統計との乖離は、あくまでも定量的には評価できない抽象的な性格のものとして、特定の社会科学の視点からみた政府統計の限界性の指摘という形で政府統計批判は専ら展開されてきた。また、社会科学方法論の立場から主体的に統計作成過程が提案される場合にも、実務経験の薄弱さの故に、提案内容は理念的な性格の強いものであった。その意味では、政府統計の第二次利用における統計利用の第一段階として提起された社会科学方法論に基づく統計の吟味批判は、それが政府統計に対する何らかの形での誤差評価のような定量的な評価基準を提示するような性格のものではなく、どちらかといえば定性的な利用上の制約の指摘に終始するものであった。その意味で、このような統計批判の観点を厳密に追求すればそれだけ、利用上の制約がむしろクローズアップされるという統計利用にとって逆説的な矛盾に直面することになる。

このような立場から統計利用を展開する場合、様々な制約を指摘した上で極めて限定的な意義を持つものという条件付きで、既存の公表数字の範囲内で可能な組み替え集計さらには近似概念による推論という形で統計利用は展開されざるをえなかった。政府統計の批判的利用という二次的利用は、調査項目の回答選択肢の許容範囲で提示可能な対案の総体と比較すれば、既存集計の範囲内での組み替え集計という極めて限定的なものであった。

統計をとりまく時代的、技術的さらには制度的制約がこれまでのわが国の統計の二次利用のレベルを規定してきた。つまり、蜷川がわが国で初めて明確に二次的利用の方法論を定式化した時代、政府統計の企画設計、さらには集計過程に政府統計の作成者以外の者が主体的に関与しうる状態には全くなく、また戦後、民主的統計制度として再建されたわが国の政府統計においても、統計作成のために収集された統計原単位情報の集計処理工程は、これまで情報処理技術の面でも政府統計作成機関だけが独占的にそれを遂行しうるものであった。その後の情報処理技術の発展とその広範な普及が、大量データの処理を政府統計機関の専管業務から少なくとも技術的には解き放つことになったとの認識を当事者が持つようになったのは比較的最近のことである。

統計データが基本的に集計結果表の形で公表、提供されてきたことから、利用者にとってもこの集計結果表が統計利用の出発点であった。このことがまた、統計利用の在り方を大きく規定してきた。一方で集計表ベースでの統計利用は、基本的に利用目的に応じた詳細分類の組み替えとしておこなわれ、その方法論については、いわゆる「グループ分け論」

として展開された。他方で回帰分析等の統計的手法がクロスセクションデータさらには時系列データに適用される場合にも、政府統計については個体データではなく、一般に集計値が用いられてきた。

### 3. ミクロデータと統計の第二次的利用

ミクロデータが PUF (Public Use Files) としてすでに 30 年以上にわたり公開提供されている欧米の多くの国では、ミクロデータからの独自集計表の作成という製表利用だけでなく、統計解析手法を直接的に適用することで幾多の学術的知見、政策的な含意を持つ発見がなされてきた。それと同時並行的に、ミクロデータの統計解析的利用に固有の問題点とその解決手段についても多くの業績が積み重ねられてきている。

欧米でその利用がすでに定着しているミクロデータへの統計解析手法の直接適用については、心理学や社会学分野で実施された学術的調査結果の解析では比較的多く見られるが、政府統計を用いた分析事例が一定の広がりをもつようになったのは比較的最近である。これにはこれまでのわが国におけるミクロデータの提供の在り方が深く関係している。

指定統計を中心にした政府統計の個票ベースでの主として学術面での利用については、旧「農家経済調査」や「農林業センサス」データが農業研究者に、また「人口動態統計」の死亡データが医学・疫学研究者による利用事例がいくつか認められる。これらは、当該省庁が専門家にミクロデータを提供することで、学術面だけでなく政策に直結する知見を得ることを目的にしたものと考えられる。

調査個票の目的外使用については、使用目的に「公益性」が必須の要件として求められる。この点で研究者による学術研究については、それが政府の政策立案等に直結したものでない限り、公益性要件が充足されたとは認められない。その点で、海外で開発された統計解析手法をわが国のデータに適用することで国際比較分析を行うという形での研究目的に対しては「公益」とは認められず、目的外申請に基づく提供の対象はされてこなかった。こういったデータの利用制約の下で、わが国ではこの種の研究の多くは、省庁が各種の『白書』作成さらには中長期計画策定を目的として申請した個票データの包括承認による分析プロジェクトに学識研究メンバーとして参画する形でこれまで行われてきた。

ミクロベースでの統計利用経験を有し、個人の就業や家計の消費活動に関わる行動分析、企業の様々な特性分析といった分野で実証分析を行う研究者の層が拡大するにつれて、調査結果の伝統的な表示形式に対し、ミクロデータが持つ豊富な情報がむしろ集計することで失われるという認識が次第に広がりつつある。

## V. むすびーマイクロデータの利用可能性と利用上の留意点

本稿では、まず調査結果の提供形式という観点からわが国における統計調査の歴史的展開を概観し、集計量として提供される結果表と非集計データであるマイクロデータの違いを統計利用の在り方と関連づけて論じてきた。

そこからいくつかのことが明らかにされてきた。まずわが国の黎明期の政府統計調査を特徴づける表式調査については、それが多くの点で業務統計と性格を共有し、また統計の作成過程そのものが目標として設定されている集計表の形式を明示的に想定したものとなっていた。このため調査客体から提供を求める統計原単位情報についても、統計作成の第一段階ですでに集計量として統合されていた。このことは、調査の企画段階で想定されていた集計結果表の獲得以外の統計作成目的には転用は一切不可能ということになる。

わが国の政府調査統計は、個票の導入により新たな段階に入る。しかし、個票によって収集された膨大な統計原単位情報も、一方では集計機械による集計処理能力の制約から、また他方では印刷媒体による結果表の公表に関わる予算面での制約から現実には予め公示された集計表に限定した形で結果の提供はなされてきた。情報処理技術の進歩によりかつてよりは多様な集計表がしかも早期公表が可能になってきてはいるが、提供される集計結果表は統計原単位情報が保有する情報のうちの一部に過ぎず、それが有する潜在的利用可能性を必ずしも尽くしたものはなっていない。

すでに諸外国の多くの国で定着している公的マイクロデータ（PUF）は、調査から得られた統計原単位情報に識別子の削除その他の匿名化措置を施したもので、単独の調査から得られるクロスセクション型のマイクロデータからも、新たな視点からの独自集計さらには回帰分析等マイクロデータであればこそ可能な統計解析的利用から、公表された既存集計だけからは決して得られなかった数多くの新たな fact findings が得られている。それは学問の進歩に寄与しているだけでなく、政策効果の事前、事後評価に有効な実践的意義も持つものである。

マイクロデータが統計調査から作り上げられたいわば公共財として広く一般の利用に供されていることは、次のような二重の意義を持つ。まず第一に、広範な利用者がマイクロデータへのアクセスが制度的に保証されていることから、これまでには考えられなかったような多面的な視角からの分析が試みられる結果、データの活用範囲が飛躍的に拡大することが期待される。第二にマイクロデータが公開されることから、この種のデータを用いた新たな実証仮説について検証可能性が保証され、そのことが理論のさらなる深化を促進することとなる。その点でマイクロデータの提供に求められる公共財としての性格は、単なるデータデバインドの解消以上の意味を持つ。

さらにマイクロデータは、異なる種類の統計調査から得られた統計原単位情報とのリンクによって、使用可能な情報量を飛躍的に拡大することができる。統計の調査環境の悪化、さらには使用可能な統計予算の制約から、新たな調査の企画はますます困難な状況にある。

その点でも、追加的な報告負担を調査客体に強いることなく、必要な統計情報の獲得ができるリンケージの意義は大きい。

マイクロデータは、同一個体に関する統計原単位情報を相互に縦断的にリンクすることで、パネル分析というこれまでの統計データがなしえなかった新たな解析可能性を提供する。パネルデータを用いれば、異時点間の位相から集団を構成したいわば動学的類型という時間の要素を集団構成に内在化し、このようにして構成された集団間の属性あるいはパフォーマンスを相互に比較するというこれまでの統計利用には見られない新たな分析視角を提供する。パネルデータについてはすでに多くの解析手法が実用化されており、個体差をコントロールした分析、クロスセクションデータと時系列データの乖離についてもこのような視角から接近が試みられている。

最後に、マイクロデータの利用者に科せられる倫理的要請について指摘することで、本稿のむすびとしたい。マイクロデータについては、種々の匿名化措置が施されているとはいえ、集計データに比べ個体識別の可能性が大きい。統計調査の場において調査客体は個人あるいは組織の個体情報が露見しないと的前提の下に統計原単位情報を提供している。この点についての信頼関係の上にそもそも統計調査は成り立っている。「悪意の利用者」によるマイクロデータの不適切な使用は、統計調査そのものの存在を危うくするものとなる。このような不当な使用を防止するための法制度の整備に加え、統計の利用者自身の倫理性が問われることになる。逆に、「悪意なき利用者」による有効な fact findings を通じた現実認識の豊富化、調査過程へのフィードバックを通して統計はより内容充実したものとなり、その社会的意義も拡大するものと考えられる。

(注1)「物産表」の報告対象となった農産物、工業製品は、以下の通り。

米、雑穀、塩類、油類、茶葉、砂糖、藁類、紙類、漆、蠟、木綿、麻紵、蚕印紙、生糸、羽毛皮、革類、海産、魚草類、漆器、鉄器、磁器、織物、金石類、染草、木材、薬品、雑品

(注2) 1883年の「農商務通信規則」によって、初めて全国統一の報告様式が採用された。

(注3) Title 13. United States Code Census

(注4) 2004年1月現在、ヨーロッパの21カ国のアーカイブがCESSDAに加盟。またCESSDAは、ヨーロッパ以外では、オーストラリアやニュージーランドなど5カ国のアーカイブと、さらにアメリカ(9大学)、カナダ(4大学)に設置されたデータセンターとリンクされている。

<http://www.nsd.uib.no/Cessda>参照

(注5) マイクロデータ提供の見返りとして科学研究機構は助成金を中央統計局に提供している。また同機構は、マイクロデータファイルの改善等の目的のために、提供者と利用者の協議機関として、1994年に科学統計本部(Scientific Statistical Agency: WSA)を設置した〔(9)p. 11〕。

## 参考文献

- (1)日本統計研究所編『日本統計発達史』1960年
- (2)鮫島龍行・相原茂『統計日本経済』筑摩書房、1971年
- (3)松田道夫『統計法の解説』（三訂版）（財）全国統計協会連合会 1973年
- (4)総務庁統計局統計基準部監修『統計行政の中・長期構想』1985年
- (5)総務庁統計局統計基準部監修『統計行政の新中・長期構想』1995年
- (6)菅宜紀「諸外国における標本データの提供状況」『統計情報』1996年5月
- (7)「ドイツにおけるマイクロ統計データの匿名化の条件と開示状況」『A02 マイクロデータ利用の社会制度上の問題』資料 No.4、1997年
- (8)浜砂敬郎「ドイツ連邦統計局によるマイクロデータの提供」『統計』日本統計協会、1998年8月号
- (9)山家善行「諸外国における標本データの提供状況（1）」『統計情報』1998年12月
- (10)甲賀智子・小松聖「諸外国における標本データの提供状況（2）」『統計情報』1999年1月
- (11)石田 晃「アメリカ、カナダにおけるマイクロデータの現状について」『マイクロ統計データの現状と展望』（法政大学日本統計研究所『研究所報』No.25）1999年 所収
- (12)石田 晃「オーストラリア、ニュージーランドにおけるマイクロデータの現状について」『マイクロ統計データの現状と展望』（法政大学日本統計研究所『研究所報』No.25）1999年 所収
- (13)浜砂敬郎「ドイツ連邦統計法におけるマイクロデータ規定と匿名化措置」『マイクロ統計データの現状と展望』（法政大学日本統計研究所『研究所報』No.25）1999年 所収
- (14)森 博美「イギリスにおけるセンサスマイクロデータの提供」『マイクロ統計データの現状と展望』（法政大学日本統計研究所『研究所報』No.25）1999年 所収
- (15)松田芳郎『マイクロ統計データの描く社会経済像』日本評論社 1999年
- (16)松田芳郎・美添泰人・伴金美編『マイクロ統計の集計解析と技法』（講座マイクロ統計分析2）日本評論社 2000年
- (17)北沢良継(2001)「パネルデータ計量経済学の最近の動向」『エコノミックス』（九州産業大学経済学会）Vol.6,No.1,pp.89-99
- (18)総務省統計局統計基準部国際統計課編『諸外国における統計の制度と運営』その24第2分冊、2002年8月
- (19) 総務省統計局統計基準部国際統計課編『諸外国における統計の制度と運営』その24第3分冊、2002年8月
- (20)各府省統計主管部局長等会議編『統計行政の新たな展開方向』2003年
- (21)北村行伸(2003)「パネルデータ分析の新展開」『経済研究』Vol.54,No.1
- (22)Dancan,G.T., et al, Private Lives and Public Policies, National Academy Press, Washington D.C. 1993

(23)Citteur,C.A.W.,Willenborg,L.C.R.J.,Public Use Microdata Files: Current Practices at national Statistical Bureaus, Journal of Official Statistics, Vol.9,No.4,1993

(24)Dale,Angela,The Decennial Census of Population: Do we still need one? Manchester Statistical Society. 1995