



MICROECONOMETRICS

マイクロ統計データの 調査・開示・利用・保存

北村行伸 Kitamura Yukinobu

1956年生まれ。慶應義塾大学経済学部卒業、ペンシルバニア大学大学院修士卒、オックスフォード大学大学院博士卒。OECD エコノミスト、日本銀行金融研究所研究員、慶應義塾大学商学部客員助教授を経て、現在、一橋大学経済研究所教授。経済学博士。著書に「パネルデータ分析」（岩波書店）ほか。専門分野：応用計量経済学、マクロ経済学、金融・財政論、公共経済学。

連載の第2回はマイクロ統計データの調査とそれを公開し利用するための仕組みについて解説する。1990年代以後ようやく政府マイクロ統計の開示が徐々に行われるようになり、それに応じてマイクロ統計を使った実証研究が蓄積されるようになってきた。しかし、その研究は質量ともにまだ不十分であり、今後一層のデータ利用の促進が望まれるところである。

1 はじめに

およそすべての統計は何らかの目的にもとづいて収集されたものである。統計データが継続的に収集・利用されているということは、国民経済生活その他の状況を示す重要な指標が蓄積されていることを意味している。国家あるいは政府が国民の生活・経済実態を知らないで善政が行われるということはあり得ない。また、国民の実態がわかってはじめて何を改善し、何が不必要かということも明らかになる。たとえ政策の関心が国民全体のマクロ経済の改善にあったとしても、その基礎になるデータはマイクロ統計である。実際のマクロ統計はマイクロ統計を積み重ねて作られた加工統計

であり、マクロ統計というものが初めから観察されるわけではない。

政府マイクロ統計はそれぞれの省庁の政策企画・立案のための基礎資料として集められたものが多く、マイクロ経済主体の行動を全般的に分析する目的で調査されたものは少ない。また逆に、各省庁が似たような質問を同一の経済主体（主として企業）に対して繰り返すなど質問される側への負担も過剰になっている。このことは近年、マイクロ統計調査が次々に行われるようになるにつれ様々な問題を生じさせている。

そのようななかで、内閣府では2004年より、「経済社会統計整備推進委員会」が設置され、その報告書『政府統計の構造改革に向けて』を受けて、設置された「統計制度改革検討委員会」では吉川洋教授（東京大学）を委員長に統計行政に関する新しい考え方が議論されており、21世紀に向けた統計行政のあり方や統計利用促進について従来よりさらに踏み込んだ議論が行われている。ここでは、統計の整備は日本再建の基礎事業中の基礎事業であるという認識の下に、①公共財としての統計、②加工統計を含む統計体系の整備、③統計の司令塔機能の強化、④法制度の見直しを含む

取り組み、の4点を軸に議論がなされている¹⁾。

2 ミクロ統計の基本デザイン

読者のほとんどは統計の利用者であり、ミクロ統計をどのように集めるかといった問題には関心が薄いかもしれない。しかし、統計データを使うすべての研究者は、自分が使う統計の収集目的と標本抽出の仕方、質問票の内容を熟知した上で、その統計が自分の研究に適しているかどうかを判断することが望まれる²⁾。

卒論や修論で、利用しているデータと本当に検討したい問題の間に大きなギャップがあるケースを多々見受ける。その原因はミクロ統計がほとんど学生に公開されていないということもあるが³⁾、実証研究に入る前に自分の使うデータを十分に吟味していないことが大きいように思う。

政府統計に限定して言えば、各省庁の統計は、統計法に従って、収集の目的、期間、方法、公開形態などが明示されており、その目的を実施するための予算を確保し、それに応じて標本数などが決定されている。したがって、各省庁が政策策定のためとはいえ、既存の統計枠組みから大きく逸脱することは許されない。容易に想像できるように、統計の設置時期が古くなるほど、現代的な分析枠組みから必要とされる情報とマッチしなくなることが多くなる。もちろん、質問票の内容に変更を加えることで追加的な情報を持つことはできるが、基本的枠組みを変えることはできないので改善には限界がある。

したがって、新規に統計調査を行うにあたっては、その統計収集の目的が、ある程度普遍的なものであり、収集方法や期間についても長期にわたって継続しても問題がないものを選択する必要がある。視点を変えて、いい統計とはどういうものかということを考えてみよう。先に触れたように現在、「統計制度改革検討委員会」では政府ミクロ統計の公共財的性格が強調されているが、事後的にできた統計をその出来不出来にかかわらず公共財であるというだけでは不十分であろう。そもそも政府ミクロ統計は特定の統計利用者のためだけにデザインされるべきではなく、長期にわたって不特定多数の利用者の利用に供すべきものである⁴⁾。もちろん、統計の性質にもよるが、時々の

思いつきで新たな質問項目を加え、その結果、継続して調査すべき項目が削除されるといったことが起こると、短期的に何らかの事実は知ることができても、長期的にみれば、大きな損失を生じさせることになるということがありうる。質問票の作成に関与する立場の方には、この点を特に注意して審議していただきたいと思う。

統計調査の実施には、調査の企画、指導、実地調査、集計、公表の段階がある。実地調査には調査票の配布、記入依頼、指導、回収などの業務が含まれている⁵⁾。このような調査を実地に行う主体は、個人で行うこともあるし、民間調査会社を実施することもある。しかし、大規模な政府統計調査では、都道府県レベルあるいは市町村レベルの地方自治体に委託されることが多い。言うまでもないが、このような統計調査がつつがなく実施されるためには、信頼でき、十分な研修を受けた調査員が全国にいて、一定の期間内に調査票を配布し回収されることが前提になる。このような統計制度のインフラストラクチャーが整備されていること自体が極めて貴重な資産であり、我が国の場合、それは十分に整備されていると考えてよいだろう。

具体的な調査票の設計については、対象分野の専門知識の他に、統計学や心理学の基本知識が必要となる。調査票の設計が間違っているために混乱が生じ、それが統計分析の誤差となって表れるようなことは回避しなければならない。清水(2000, pp.19-20)によれば、次のような点に注意する必要がある。

①不必要な質問は含めないこと。②質問を設定する前に、それから作成される統計を想定すること。③質問や説明は平易な言葉を用いること。④回答に難しい表現を求めないこと。⑤回答を誘導するように質問しないこと。⑥複数の内容を持った質問をしないこと。⑦個々の質問に対する無回答をすくなくすること。⑧質問量は過度にならないようにすること。⑨質問の配置は前問の影響を受けることのないようにすること。⑩対象者が回答しやすいような配列にすること。⑪主題の転換部では話題の変更が明確にわかるようにすること。

著者の経験でも、『慶應家計パネル調査』の調査票作成にオブザーバーとして参加したり、家計経済研究所の『消費生活に関するパネル調査』の

調査票配布回収時の苦勞話をたくさん聞いており、調査票の作成には必要以上に神経を使わなければ、思ったような結果が得られないことがわかっている。

実際に、海外の統計調査で調べられており、それを元に興味深い結果が出ている研究があり、我が国でも同様の実証研究を行ってみたいと思っても調査自体が拒絶される場合がある。すなわち、すべての社会にはそれぞれ社会的なタブーや政治的に調査することが不適切と判断される項目がある。具体例として、アメリカのデータに人種に関する指標やIQに関する指標が含まれているのに対して、日本の社会調査のデータには戸籍に関する情報やIQに関する情報は一般に収集されていないために、アメリカとの比較研究が難しい場合がある。逆に、アメリカでは家計単位で消費に関する家計簿を3-6ヶ月間も毎日綿密に記入してもらうことは不可能であるが、日本では、この統計調査（『家計調査』）は長期に維持継続されてきている⁶⁾。

調査票の内容を小規模なサンプルに対して実験的に配布し、回答の難易度や拒否率を検証して、さらに調査票を改善し、最終的に本格的な統計調査を行えるまでに準備を整えておくのが一般的である。

3 標本調査の考え方

さて調査票ができて、あとはその調査票を配布すればいい段階にきている。では、これを誰に配布すればよいのだろうか。誰に配布すべきであるという指針はあるのだろうか。当然考えられるのは、できるだけ適切に社会全体の分布に近い形で配布したいが、全国民に対して調査するほどの費用はないという現実である。この問題に対して考えられたのが標本抽出の理論である。標本抽出の考え方を経済学的に言い換えると「社会全体（母集団）をできるだけ正確に、しかもできるだけ費用を少なく知るにはどうすればいいのか」ということである。それぞれの事象に対してどれぐらいの標本を集めれば信頼のおける統計情報が集められるかということはきわめて重要な問題である⁷⁾。

第1の問題は、母集団をすべて記入した台帳（名簿）を手に入れるということである。世論調

査であれば、選挙人名簿や住民基本台帳などが使われるであろうし、もう少し雑なものでよければ、電話帳や会員名簿、同窓会名簿などが使われることもある。標本規模が小さな調査であれば、これらの台帳（名簿）からできるだけランダムに標本を抽出する方法を考えればいい。標本規模が大きい調査では第1段階で「市町村」あるいは国勢調査の「調査区」を抽出し、第2段階で個人あるいは世帯を抽出する⁸⁾。標本規模に応じて抽出段階を3段階以上にすることもあり、これをまとめて「層別（層化）多段抽出法」と呼ぶ。

層別2段抽出法を使ったとして、第1段階の抽出方法は、市町村の大きさ（人口規模）に基づいて、規模の大きさに比例した抽出比率による確率比例抽出法を利用する。すなわち、全市町村の全人口に通し番号をつけ、乱数表によって選んだ数字が該当する個人のある市町村をまず選択するということである。第2段階での個人あるいは世帯の抽出方法は①単純ランダム抽出、②系統抽出⁹⁾が考えられるが、通常②が用いられる。抽出に用いる台帳は③住民基本台帳か④選挙人名簿である¹⁰⁾。

では、このようにして選択した標本は母集団の統計特性に対してどれぐらいのバイアスを持っているのだろうか。たとえば、調べたい項目が年齢や家族構成であれば、全数調査である国勢調査の年齢や家族構成の分布を母集団として、それからの乖離を調べればいい。層別2段抽出法の場合、標本抽出誤差の大きさは第1次抽出単位の散らばりの平均的な大きさに関連していることが知られている。調査前に、そのバイアスを知ることはできないが、同種の調査でどのような項目でバイアスが大きいかわかることはできる。一般に層別化をうまく行えば、誤差分散を減少させることが可能になる。それを達成するためには層内分散を減少させ、層間の分散を大きくするように層別化を行えばよいとされている。このことの直感的な考え方は次のようなものである。母分散の大きな母集団の平均値を調べるには多くの標本を必要とするが、母分散の小さな母集団の場合には少ない標本数でも同じ精度が保てることを意味している。とすれば、分散の小さな層ごとに無作為抽出を行って標本平均の分散を小さくすることで、できるだけ小さな標本数で精度の高い統計量を得ること

ができる、あるいは同じ標本数であれば分散を減少させることができるということである。

これらの結果は、地域と関連の深い項目について調査する場合には層別を慎重に行う必要があること、調査地点数と層サンプル数とのバランスを十分考慮すべきことを示唆している¹¹⁾。

さて、このようにして確定した標本に対して調査を実施したとしても問題はまだ残っている。まず、調査員が対象個人あるいは世帯に調査への協力を依頼しても、必ずしも引き受けてくれるとは限らない。近年、ライフスタイルが多様化し、調査員が訪問しても不在であったり、個人情報保護の意識が高まり、たとえ政府統計であったとしても、なかなか理解が得られずに調査を拒否する個人あるいは世帯が増えている。その場合、同一層内で標本を再抽出する必要が出てくるが、新たに系統抽出を行えば、事前に行った抽出の無作為性が崩れることになる。この場合、事前に拒否率も織り込んで多めに標本抽出しておくことが望ましい¹²⁾。さらに、得られた調査票を点検して、誤記や見落としがあった場合には、単純なミスであれば調査側で修正し、対象者に聞かなければ分からない項目に関しては対象者を再訪して、調査票を埋めたり修正したりする必要がある。このような過程を通して貴重な統計データが集められてくるのである。

統計の利用者には、統計を収集する側の苦労はなかなか見えてこないが、単に収集費用には還元できない陰の働きがたくさんあることを知っていただきたい。ある研究者は、調査員について対象者を訪問し、実際に調査を依頼することの難しさを経験してから、結果として得られた調査票の情報はあだやおろそかにはできないことを実感したと述べていたが、著者もまったく同感である。

実際の統計調査では標本抽出の理論にもとづいて、できるだけ少ない標本で、できるだけバイアスの少ない統計量を得ようという努力がなされているのである。このような標本調査あるいはその理論的基礎である標本理論はマイクロ計量経済学の直接の研究対象ではないが、きわめて密接に関連した分野である。繰り返しになるが、自分が使っているマイクロ統計データがどのように設計され、どのような過程を経て標本抽出され、どのように取りまとめられたのかを知ることは大切なことで

ある。また、同時に、調査側も、調査の経緯について詳細に報告する義務がある。

4 ミクロ統計調査の開示と保存

収集された調査票は、取りまとめられデジタル・データとして入力され、それを統計的に集計したものが報告書として公表されるのが普通である。また、政府統計であればそれをを用いて政策形成に定期的に利用されることも多いし、国民経済計算体系(SNA)の基礎統計として利用されることもある。

たとえば、総務省の発表する消費者物価指数(CPI)の作成には「小売物価統計調査」や「家計調査」の情報が用いられている。日本銀行の「全国企業短期経済観測調査(企業短観)」の業況判断は必ず新聞のヘッドラインを飾るほど注目度が高い統計であり、民間エコノミストも短観の結果を利用して景気分析を行っている。しかし、政府統計のなかには、報告書は出ているが、それがどのような政策に利用されているのか、政策形成以外の目的でどのように使えばいいのかが不明なものもあるように思われる。また、政府関係者以外の研究者がまったく利用したことのない政府統計も多くある。

先に紹介した内閣府が昨年6月に公表した『政府統計の構造改革に向けて』のなかでも、統計情報の多様かつ高度な利用の促進がうたわれているが、従来行われてきた、公表された報告書や政府がインターネット上で公開した統計を利用する方法、あるいは、総務大臣の承認を経て調査票の目的外利用(統計法第15条、第15条の2に基づく)の枠組みのなかで利用する方法、はとみに限界があり、今後の見直しが必要である¹³⁾。

先に紹介したように政府統計は公共財であるという認識が強調されるようになってきた。この考え方には、公共財が活発に利用されるべきであるという側面がある。しかし、政府統計の利用は限定的であり、客観的に見て決して公共財と呼べるような状況にはなく、あえて表現すればクラブ財とも呼ぶべきではないだろうか。実際、政府統計を利用してきたのは、政府か政府系機関の主催する研究会や実務に参加した研究者か、文部科学省科学研究費の特定領域研究(「統計情報活用の

フロンティアの拡大—マイクロデータによる社会構造解析—」平成8-10年、領域代表松田芳郎教授)に参加して政府統計の目的外利用申請の負担が大幅に軽減された研究者か、一橋大学経済研究所附属社会科学統計情報研究センター内にこの科学研究費の後継として新たに設置された、マイクロ分析セクションを通して公開しているリサンプルデータを利用するか¹⁴⁾、あるいは独自に総務省統計局に対して政府統計の目的外利用申請を行うかなどの道があるが、すべての研究者が政府統計を気楽に利用できる状況にないことは確かである。

また、多大な費用をかけて収集した統計を一定の保存期間後に処分したり、あるいは積極的に処分しなくとも、維持管理を行わなければ原資料およびデジタル情報が劣化して、使い物にならなくなるようでは、これまた公共財とは呼べない。データの保存については、その統計の歴史的意義を認めて保存すべきであるという考え方の他、将来世代が新たな視点、新たな技術を持って過去のデータを分析する利用権を保護すべきであるという考え方もできる。

政府統計はその時々 of 社会の実態を反映する貴重な資料であり、適切な統計調査に基づいて適切な質問票が作られている限り、その公開を拒む理由もなければ、一定期間後に処分する理由も見当たらない。とすれば、政府統計を維持保存しておくアーカイブ機能を充実させることも重要な課題であることがわかるだろう。具体的には、統計局のなかに政府統計アーカイブを作るか¹⁵⁾、公共機関、たとえば、国立大学法人のなかでアーカイブ機能を担う用意のある大学が責任を持って管理するという事も考えられる¹⁶⁾。

情報公開法が導入されて公開できないと判断されるような資料がアーカイブに入ることなく処分されていると伝聞するが、相当のコストをかけて収集した政府統計はすべて長期にわたり保存し、公開をすすめることが必要である。政府統計は、その時々 of 社会の実態を伝える貴重なドキュメントであり、逆に、調査対象者の機密保護は大前提であるが、その資料全体が機密性があったり、公開前に処分される類の政治的意味があるとは考えにくい。現代の学者がツタンカーメン王の死因を分析できるようになったように、数千年後の学者の手にゆだねれば、我々には思い及びもしない方

法で平成バブルやデフレの原因を解明してくれるかもしれないではないか。

5 ミクロ統計利用者からの要望

これまでの議論からも明らかなようにマイクロ統計データ、とりわけ政府統計は、一部の研究者にしか利用できなかった。幸いにも、著者の所属する一橋大学経済研究所附属社会科学統計情報研究センターでは政府マイクロ統計データの積極的な利用促進を進めており、各種の政府統計を利用させてもらうことができた。

マイクロ統計データを学問研究上利用することによって、(1)社会経済のマイクロ構造の理解が深まる、(2)政策分析の質が向上する、(3)統計の問題を発見することによって統計の質の改善が期待される。

(1)(2)の点については一般に認識されているが、(3)については、これまで利用が制限されてきたこともあってあまり論じられてはこなかった。繰り返し述べているように、政府統計を公共財と見なすのであれば、公共財の利用者の声を反映させて、公共財自体の質を高めることも考えなければならない。

細かいミスでも気づいたときに調査側に報告し、調査側がそれに真摯に対応して、少しずつ問題を改善していけば次第に完全な統計データに近づくと考えればいいのではないだろうか¹⁷⁾。各省庁の統計担当部署では、統計に関する質問や問題の指摘を受け付ける窓口を用意すべきであろう。担当者に直接苦情が届いた場合、苦情自体が握りつぶされる可能性もあり、また、握りつぶさずとも放置されて忘れ去られる危険性もある。

データの中には未記入の部分もあり、また調査側も気づかない誤記も残っていることが間々ある。サブ項目を足し合わせれば項目合計になるはずが一致していない場合もある(たとえば、消費や所得の収支合計が個別部分を足し合わせたものと一致しない場合)。調査票のある部分の回答に対しては欠損値は(-)で表記してあるのに、他の場所では欠損値が(0)と表記されていたり、(9999)と表記されていたりすることもある。もちろん、始めから完全なデータを求めることはないが、少なくとも、研究者がかなり注意深くデー

タを見ないと見過ごしてしまうような混乱は避け
てほしい。あるいは、少なくとも、誰かが気づい
て、問題点を指摘すれば、以後の利用者に配布さ
れるマイクロ統計データではその問題点が改善され
ていることが望ましい¹⁸⁾。

これまで、マイクロ統計データの利用許可が出る
と、各省庁によって様々な形の磁気データが貸与
される。通常、そこにはテキスト形式で無味乾燥
に打ち込まれた数字が羅列されているだけである。
それをデータセットとして読み取るためには、
SASなどのソフトウェアでプログラムを書いて、
何行目から何行目は変数A、何行目から何行目は
変数Bという風に数字に意味を持たせる必要がある¹⁹⁾。
このような、読み取り作業を経て始めて、
実証研究に使える形になるのである。欧米での政
府統計の公開事例を見ると、データは利用者に渡
される時点でSTATA, SAS, SPSSなど主要なソ
フトウェアで読み込めるデータフォーマットにし
てあり、利用者は自分の欲しいデータフォーマ
ットを指定できるようになっている。これも、先
のデータ上のミスと同じで、一度各ソフトのフォー
マットに直してしまえば、あとはそれをコピーす
ればいいだけなのでそれほど手間ではないはずで
ある。このような対応も、統計が公共財である
とか、統計サービスも省庁の業務であるという意識
が出てくれば、自然に発生してくるものと期待し
ている。

もう一つの要望として、政府統計に1-2の追
加的質問を加えることで、利用者の作業の
潜在的な誤りが大幅に削減できる場合がある
ことを指摘しておきたい。たとえば、総務省の「全
国消費実態調査」では年間収入のデータは入
っているが、年間に支払った税、社会保
障保険料が調査されていないために、年
間の可処分所得を求めるために、税と
社会保障保険料を家計属性から家計
毎に推定して、それにもとづいて可
処分所得を求めている。しかし、多
くの家計は年間所得を記入する際
に納税データを利用しているとす
れば、税や社会保障保険料も同
時に記入してもらうことができ
れば、このような手間が省ける
だけでなく、統計データに反
映されていない要因によって、
税控除を受けていることもある
ので、そのような測定誤差も
回避できる。これは、このデ
ータの目的外利用者の利用
目的が調査票作成者に伝わっ

ないために生じている問題だが、
実際、このデータを利用する
ものの大半がこの問題に悩ま
されていると思われる。

最後に要望するだけでなく
マイクロ統計データの利用者
からの貢献についても触れて
おきたい。それは政府統計
アーカイブだけでなく、研
究者がマイクロ統計データ
を利用することで得た重要
な推定パラメータ（たと
えば、財別の需要の価格
弾力性や通貨需要関数の
利子弾力性等）や研究の
過程で作成した重要な加
工データ（たとえば、企
業別、産業別のトービ
ンの q 、資本ストックの
市場価値、企業別、産
業別の生産性指標、産
業別デフレタ等）など
も個人ホームページで
公開したり、研究を
支援した研究機関の
ホームページなどで
公開されることが望
ましいということである²⁰⁾。
また、それらのデ
ータがどこでどの
ような形で公開さ
れているかをとり
まとめてリンクを
張り、常にア
ップデートして
いるウェブサイト
が作られること
も望ましいだ
ろう。

6 おわりに

今回は連載の第2回ということで、
マイクロ計量経済学の資料にあ
たるマイクロ統計データの作成、
公開、利用、保存について論
じてきた。なかなかマイクロ計
量経済学の本題に入らないの
でもどかしく思っている方も
いるかもしれないが、マイクロ
計量経済学は、研究対象にな
る経済主体に関する統計が
いかなる状況で調査され、
いかなる過程を経て集計され
てきたかを知ることから始
まるものだということがわか
っていただけたのではない
かと思う。また、マイクロ統
計の利用は経済分析の多様
性を保障するだけでなく、
政策分析の質の向上を通
して、ひいては政府の政策
立案能力の向上にもつなが
ること、さらに、統計利用
者の声を統計に反映させ、
統計の質の向上に結びつけ
ることも重要であることも
強調しておきたい。

冒頭で指摘したとおり、
統計制度は大改革を迎
えようとしている。政府
統計を含め、マイクロ
統計が、より開かれ
た、より機動的な
ものになっ
ていくものと
期待している。

注

1) 詳細については内閣府のホームページや舟岡（2005）、
竹内（2005）を参照されたい。

- 2) これは新しい機械を使う前のある程度、操作マニュアルを読んでおくのが望ましいことに似ている。機械の場合には、細かい操作は必要に応じてマニュアルを見直す、統計データに関しても、必要に応じて質問票に立ち戻ることが肝心である。
- 3) 政府マイクロ統計は一般に学生には公開されていないので、学生が使えるデータは民間ベースで集められた統計(たとえば、日経ニーズ・データ)や家計経済研究所の「消費生活に関するパネル調査」、政策投資銀行の企業財務データなどに限られている。
- 4) ここで、将来にわたるすべての利用方法を想定して、統計調査を行うべきだと主張するつもりはないし、それは実際問題として不可能であろう。しかし、共有地として多目的に使える広場に、その時の勢いでスケートリンクを作ったりすると、すべての人の利用に供する公共財としての将来価値は低下すると考えられるのではないだろうか。統計調査もできるだけ普遍的な質問事項を中心に設計し、時間的な変化に対して根本的な変更を必要としない構造を求めべきであろう。
- 5) 統計調査方法については、鈴木・高橋(1998)、豊田(1998)、清水(2000)を参照。
- 6) もちろん、1回限りの調査であれば、タブーにチャレンジすることも可能かもしれない。すなわち、クロスセクション・データであれば、新たなサンプルを選ぶことができる。しかし、パネルデータの場合、タブーにチャレンジすることで、サンプルから大幅に脱落してしまうと取り返しがつかないというリスクがあり、慎重にならざるを得ない。
- 7) フジテレビの「トリビアの泉」で青山学院大学の美添泰人教授が、それぞれの事象に関して、必要な標本数を概算されているのは、標本理論の成果にもとづいたものである。標本理論の詳細については、たとえば、Cochran(1977)を参照されたい。
- 8) これを層別2段階抽出法と呼ぶ。層別抽出することのメリットは、抽出にともなう費用が大幅に軽減できることにある。第1に、第1段階の抽出に必要な全国の市町村名称と人口はすぐ手にはいるので、そこで選ばれた市町村についてのみ住民基本台帳を入手すればよい。第2に、第2段階で抽出された標本は比較的狭い範囲に住んでいる(同一市内、町内)ので調査員が訪問するのも楽である。詳しくは豊田(1998、19-20章)を参照されたい。
- 9) 系統抽出とは抽出間隔 k とし、スタート番号から等間隔に標本を選んでいく方法であり、無作為抽出に代わり用いられることが多い。
- 10) 具体的な抽出の手順については鈴木・高橋(1998、第2部、第4章)を参照。
- 11) 詳細については鈴木・高橋(1998、第2部、第5章)を参照されたい。
- 12) 実際の統計調査でも回答拒否が広範に存在することは知られているが、その対処方法がどのようにとられたかはあまり明らかにはされていない。
- 13) 政府統計の開示の実態や法的、制度的側面については、松田・濱砂・森(編)(2000)を参照されたい。
- 14) 一橋大学経済研究所附属社会科学統計情報研究センターでの利用促進の実態については松井(2005)を参照されたい。
- 15) 国立公文書館のデータセクションとして保存することも考えられるのではないだろうか。

- 16) イギリスではエセックス大学が統計アーカイブの機能を担っている。
- 17) 一昨年には内閣府の国民経済計算において消費税の取り扱いが5%に引き上げられた後も3%として計算されていたことが報じられたが、このケースでは指摘を受けてからかなり長期にわたって問題を放置していたことが判明した。間違いはないに越したことはないが、誰でも気が付かないミスはあるものである。問題は気づいた時に、いかに速やかに対処するかということである。
- 18) 同じ問題点を個々の研究者が繰り返し修正を加えるということはいかにも無駄であるし、気づく人もいれば気づかぬ人もいるということでは、同じマイクロ統計データで結果が微妙に違うということが起こる。
- 19) 通常、目的外利用による政府統計の利用期間は3-6ヶ月であり、慣れていない人であれば、データを読み取るだけで使用期限が過ぎてしまうことも起こり得る。
- 20) 実際には、内閣府経済社会総合研究所ではESRI-HISTAT-JIPDATABASEとして、日本の産業別資本ストック推計や生産性推計の結果を公表しているし、一橋大学経済研究所附属社会科学統計情報研究センターでは家計調査をパネルデータとして独自に集計した結果を公表するなど、多くの研究機関が加工データを公表するようになってきている。

参考文献

- [1] 舟岡史雄(2005)「日本の経済社会統計：どう変えていくのか」、『経済セミナー』、2005年12月号、pp.20-3.
- [2] 清水誠(2000)『統計体系入門』、日本評論社。
- [3] 鈴木達三、高橋宏一(1998)『標本調査法』、朝倉書店。
- [4] 竹内啓(2005)「統計の改革のあるべき「方向」を求めて」、『経済セミナー』、2005年12月号、pp.37-40.
- [5] 豊田秀樹(1998)『調査法講義』、朝倉書店。
- [6] 松井博(2005)「統計データのより高度な利用に向けて」、『経済セミナー』、2005年12月号、pp.34-36.
- [7] 松田芳郎、濱砂敬郎、森博美(編)(2000)『講座マイクロ統計分析(1)統計調査制度とマイクロ統計の開示』、日本評論社。
- [8] Cameron, A. C. and Trivedi, P. K. (2005) *Microeconomics: Methods and Applications*, Cambridge University Press.
- [9] Cochran, Willam. (1977) *Sampling Techniques*, 3rd ed, John Wiley&Sons.
- [10] Lee, Myoung-Jae. (2005) *Micro-Econometrics for Policy, Program, and Treatment Effects*, Oxford University Press.
- [11] Winklemann, Rainer and Boes, Stefan. (2005) *Analysis of Microdata*, Springer.
- [12] Wooldridge, Jeffrey. M. (2003) *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.