

「経済統計・政府統計の数理的基礎とその応用 2011」開催にあたって

山本 拓
日本大学経済学部

本研究集会は、科学研究費プロジェクト「経済統計・政府統計の数理的基礎とその応用」（2011年度－2014年度、研究代表者：山本 拓）に基づく、2011年度の研究集会である。国友直人と山本拓がオーガナイザーをつとめた。

本プロジェクトの目的は、経済統計・政府統計における主要な課題の、技術的および制度的問題を、統計学的な立場から理論的・学術的に検討し、具体的解決策を提言することである。

経済統計、とりわけ政府統計は、経済・社会の動向を理解し、政策を実施、評価するためには不可欠な情報であることは言うまでもない。最近では **evidence-based policy** ということもよく言われ、政府統計の重要性は一般に広く認識されつつあると思われる。

しかし、経済統計・政府統計への信頼性は、近年必ずしも増しているとは言えない状況である。経済社会の急激な変化に伴い、政府統計の質の確保が困難になりつつある。マクロ経済統計の側面では、GDP 統計などに代表されるマクロ公表系列の質と信頼性の問題、信頼性の高い将来人口の推計の問題、地域による経済情勢のばらつき把握などの問題を挙げることができる。またミクロ経済データにおいては、統計調査をとりまくプライバシー意識の高まりから、調査精度の確保が難しくなりつつあるという問題や、情報開示と秘密保持の両立という匿名化問題などを挙げることができる。

これらの困難さの根本的原因の一つは、先に述べた政府統計の重要性の認識とは裏腹に、財政難から政府統計関連の予算が長期的な漸減傾向にあり、政府統計部局には残念ながら上記のような問題に対処するための十分な人的・財政的資源が確保されていないことにある。中・長期的視野に立ち、理論的あるいは学術的基盤まで考慮した新たな確固たる統計的方法を検討するには、余力に乏しい状態なのである。また政府部局の人材採用・育成メカニズム自体も基本的に旧来通り制約的であり、例えば大学院修了者を自由に雇用することはできず、日本政府の統計部局の人的構成は大学院修了者を多量に抱える欧米先進国や新興国とは異なる状況となっている。

新しい統計学的知見の導入に関しては、日本の政府統計部局が分散化されているために、これまでは、個別の担当部局あるいはその時々担当者に個別に招かれた研究者によって知見や助言が提供されることが多かった。政府統計を巡る重要な論点について、担当部局をまたいでその知見が共有されることは少なかったと思われる。またそれらの話題が広く研究者間で議論されることも少なかった。そのような意味で、経済統計・政

府統計の技術的・制度的問題点を、統計学的立場から総括的に検討していくという本研究プロジェクトは、一つの新しい方向性を示している。

また一般の統計学研究者と統計職員がオープンかつ総括的に問題を検討していくという機会は、これまであまり存在しなかったと思われる。（日本統計学会において政府統計のセッションが設けられたことは多々あると記憶するが、日本統計学会に所属している政府統計職員は限られているのが実情である。）今回の研究集会は、統計学研究者と統計調査担当者がおよそ半々の割合で、さまざまなトピックについて報告を行う構成になっている。このような機会が情報交換ならびにお互いの刺激となり、経済統計・政府統計の今後の改善の一助になることを期待する次第である。

家計調査の課題と改善に向けて

宇南山 卓*

家計調査には、その精度が低いという「標本誤差」に対する批判と、偏りがあるという「非標本誤差」に対する批判があった。標本誤差については、学術的な分析には問題ないレベルであるが、景気指標としての統計ユーザーの要求水準は満たしていなかった。高い精度を確保するには、補完的な統計を活用する必要がある。非標本誤差については、専業主婦世帯の割合が過大である、低所得者が過少であるというサンプルセレクションバイアスがあった。また、高額商品に対する支出および財産所得の受取について、過少となるバイアスがあった。さらに、これらのバイアスの複合的な結果として、貯蓄率が過大になっていた。いくつかの問題点を抱えてはいるが、家計調査は「繰り返しクロスセクション統計」としては国際的にも最高水準の統計と評価できる。将来的には、家計調査をパネルデータ化することが期待される。

1. はじめに

家計調査は「国民生活における家計収支の実態を把握し、国の経済政策・社会政策の立案のための基礎資料を提供することを目的」とした統計であり¹⁾、1946年から調査されている。家計調査の結果は、基本的には消費者物価指数作成のためのウェイト算出に利用され、それ以外にも国民経済計算における四半期別 GDP 速報の家計最終消費支出の推計や景気動向指数にも採用されており、個人消費の動向を判断する指標として活用されている。学術的な関心からは、収入・消費の水準・分布の把握、消費者行動の分析、政策の評価のための重要なデータである。

しかし、家計調査に対しては多くの批判がされており、統計としての信頼性に対する疑問が指摘されている。例えば、総合研究開発機構(2008)による「市場分析専門家の立場から見た経済統計に関するアンケート調査」(以下 NIRA アンケートと呼ぶ)によれば、国などが公表している 23 の統計のうち、家計調査の評価が最も低い。加えて、新聞など一般の媒体でも、家計調査の信頼性が問題とされたり、それに対する反論が掲載されるなど関心は高い²⁾。利用範囲の広さ・関心の高さを前提とすれば、家計調査の信頼性がより厳しく求められるのはもちろんであるが、家計調査の特性からすればやむを得ない欠点への批判も多く、本質的な議論のためには論点の整理が必要である。

本稿では、これまでの家計調査の信頼性に関する議論をまとめ、その妥当性を検討する。また、過去における改善についても概観し、残された課題を明らかにする。さらに一般的な

* 神戸大学大学院経済学研究科, 〒 657-8501 神戸市灘区六甲台町 2-1, E-mail: unayama@person.kobe-u.ac.jp

¹⁾ 総務省統計局ホームページ・家計調査の概要「1. 調査の目的」を参照。

²⁾ 例えば、日本経済新聞 1990年2月21日「経済教室：大きな偏りない家計調査」、同夕刊 1993年6月12日「エコノ探偵団：あてになるか？家計調査」などを参照。

問題として、日本における収入・消費に関する世帯調査の課題について述べる。こうした整理によって、家計調査の特性を明らかにし、適切なデータの利用や統計の改善に資することを目的とする。

家計調査に関する議論を概観してみると、最大のユーザーであり批判者でもあるのが市場の「エコノミスト」であることが分かる。NIRA アンケートの対象となった「市場分析専門家」のほとんどは、家計調査に注目している一方で結果に対し批判的である。具体的な批判のほとんどは、月次の「変動」が大きすぎることの一点に集中している。変動が大きすぎるとは、家計調査の結果が、しばしば、GDP 統計や毎月勤労統計などの景気指標と説明が困難な乖離をすることを指している。言い換えれば、家計調査の精度が低すぎるとの批判である。

実際、景気動向に注目するエコノミストの多くが各種指標に1%ポイント以下の精度を要求しているのに対し、家計調査の標準誤差率は2.0%となっている³⁾。すなわち、要求されている水準よりも精度が低いという意味では、この批判には一定の妥当性がある。しかし、これは統計の目的と利用の実態が乖離していることに起因した問題であるという意味では、不可避的な問題である。

日本の家計のほとんどが月給を受け取り、1カ月を単位とした行動をとっていることを考慮すれば、家計調査が月次で調査されていることは望ましい性質である⁴⁾。すなわち、月次で調査する目的は、必ずしも景気判断のための材料を提供することではないのである。しかし、収入・支出に関する情報が月次で利用可能であるという速報性のために、GDP の推計などにも利用され、景気指標としても注目されてしまうのである。本来の目的である「家計収支の実態を把握」するために想定されている標本誤差の水準が、景気判断の目的に要求される水準よりも大きいため、批判の対象となっているのである。

家計調査に対する批判の大部分が、こうした標本誤差に関する問題であれば、調査世帯を増加させることでより精度の高い統計を作成すればよく、原理的には解決は容易である。しかし、調査の規模の拡大のコストと精度向上の便益の比較・速報性の維持に対する要求も考慮すると、単純に調査世帯を増加させることは最適な対策ではないと考えられる。

より安定的な消費動向の指標を構築するために、家計調査の規模拡大ではなく家計調査を補完する「家計消費状況調査」が2002年から開始されている。実際に、家計消費状況調査と組み合わせることで、家計調査の景気指標としての信頼性は大幅に改善している。そもそも調査統計において一定の標本誤差は不可避であることも考慮すれば、利用者としても利

³⁾ 総務省統計局のホームページから、家計調査のページ> Q&A の Q16 を参照 (<http://www.stat.go.jp/data/kakei/qa-1.htm>)。

⁴⁾ 調査期間は、例えば、英国の Family Expenditure Survey (FES) は各世帯 2 週間のみ、米国の Consumption Expenditure Survey (CES) は家計簿の記入が 1 週間、記憶による調査が四半期ごとである。

用可能な統計の標本誤差を考慮した分析手法を開発する必要があるだろう。

一方、学術的な関心に基づく家計行動の分析などにおいては、2%程度の標本誤差は大きな問題ではない。より深刻な問題となるのは、調査の結果が構造的に「真の値」から異なるという意味での「非標本誤差」である。すでに述べたように、批判の多くは「標本誤差」に対するものであるが、本稿では「非標本誤差」が本質的な問題と考え、いくつかの問題を具体的に論じた。

家計調査に対して指摘されている非標本誤差には、大きく分けて2つのタイプがある。一つは、標本である調査世帯の属性が、母集団である日本の家計全体の属性とは異なる「サンプルセレクションバイアス」である。具体的には、公務員世帯の割合が過大である、共働き世帯の割合が過少である、低所得者・超高所得者の割合が過少であるなどの批判である。もう一つは、調査世帯の属性に問題がなくても、特定の項目に対する調査が困難であるために、計測される結果が真の値と異なる「測定誤差」である。例えば、耐久消費財や冠婚葬祭費などの高額商品や、財産所得などの結果に偏りがあることが予想されており、また貯蓄率の動向が他の統計と整合的でないことも指摘されている。

実際に、それぞれの批判についてその妥当性を検証すると、公務員世帯が多いという批判は妥当ではなく結果に与える影響も小さいと予想されたが、専業主婦世帯の割合が過大であるという指摘には妥当性があった。また、所得水準の分布に偏りがあるかについては検証は困難であるが、少なくとも低所得者に関して偏りがあることが示唆されている。項目別の偏りについては、高額商品については家計調査の結果が過少である可能性が高く、財産所得についても多くの記入漏れが発生している。貯蓄率については、先行研究で指摘されていた要因に加え、ここで検証している項目の偏りが原因で、他の統計との乖離が発生していたと考えられる。

このように、いくつかの問題点を抱えてはいるが、家計調査は「繰り返しクロスセクション統計」としては国際的にも最高水準の統計である。しかし、経済学の発展に伴い、研究者の関心が変化してきており、家計調査の構造はその変化に十分には対応できていない。特に、家計の動的な行動を描写するために同一世帯の行動を追跡した「パネルデータ」の重要性が高まっており、日本においてもその構築が期待される。さらに社会経済の変化として、近年のプライバシー意識の高まりがあり、個人情報保持に敏感な世帯が増えている。安定的な調査の継続には、何らかの対応が必要であろう。こうした課題に対応するためには、現行の家計調査の改善にとどまらず、世帯の収入・支出を世帯調査からどのように把握していくかについての体系的な検討が必要であると考えられる。

本稿の構成は以下のとおりである。第2章では、家計調査の概要について述べる。実際の調査世帯の抽出や調査方法について紹介する。第3章では、これまで指摘されている家計調査への批判をまとめ、その妥当性を検証している。また、結果の公表形式についてもこれま

で指摘されている改善点を示した。第4章は、家計調査に限らず、収入・支出に関する世帯調査全般に対する展望を示している。第5章は結論である。

2. 家計調査の概要

家計調査は、1946年に始められた「消費者価格調査」が発展したものである⁵⁾。「消費者価格調査」は都市に居住する単身世帯を除く非農林漁家世帯を対象として、家計の購買行動について、その価格・購入数量・支出金額を調査したものであった。1950年から収入についても調査を開始し、1951年からは「消費実態調査」に名称も変更された。その後さらに、調査方法と費目分類が若干改正され、1953年から「家計調査」と呼ばれるようになった。1962年には母集団地域が全国に拡大され、調査対象が168市町村・約8,000世帯となり、ほぼ現行の家計調査と同じ調査形態となった。

1981年1月に収支項目分類が大幅に改正され、消費支出の5大費目分類が10大費目分類となり、現在の統計とほぼ完全に接続可能な統計となった。さらに、かつては、農林漁家や単身世帯が調査対象とされておらず家計調査の結果が日本の家計全体を代表していないと批判されていたが、2000年には農林漁家が、2002年には単身世帯が調査対象となるように拡大されている⁶⁾。現行の家計調査は、全国約4,700万世帯の家計の中から、毎月9,000世帯が家計簿を記入する方式で月々の収入・支出を調査している。

調査世帯の選定については、その一部が毎月交代する「ローテーションパネル」方式で調査されており、各世帯の調査期間は二人以上の世帯では6ヶ月、単身世帯は3ヶ月となっている。毎月新たに調査される世帯は、層化三段抽出法によって次の三つの段階に分けて選定される。第一段階として、全国の市町村の中から、地理的位置・人口・産業の特色などを考慮して「層」と呼ばれるグループに分類し、それぞれの層から一市町村ずつ合計168市町村を抽出する。第二に、抽出された市町村から無作為に調査地区を選ぶ。最後の第3段階目として、それぞれの調査地区から乱数表を使って無作為に調査世帯を選ぶのである。言い換えれば、調査対象が日本の世帯全体の縮図となるように、できる限り無作為に抽出されているのである⁷⁾。ただし、実際に選定された世帯が調査に協力することを拒否した場合には、一定の手続きに従い代替の世帯が選ばれ調査される⁸⁾。

調査項目としては、勤労者及び無職者世帯については収入と支出を、個人営業世帯などの

5) 以下の記述は、総務省統計局ホームページにおける「家計調査のページ>家計調査の概要>家計調査の沿革」にしたがっている。戦前にも現在の家計調査に類似した世帯調査は行われていた。戦前の状況については、多田(1989)が詳しい。

6) 1999年以前は、農林漁家に対しては農家経済調査などが実施されていた。また、1995年以降単身世帯も単身世帯収支調査の調査対象であった。

7) 松田・伴・美添(2000)は標本調査論の観点から、家計調査の調査方法について厳密に論じている。

8) 家計調査は、統計法に基づく基幹統計(旧統計法に基づく指定統計)であり、調査拒否は認められていない(統計法第13条)。しかし、現実には調査を強制することは困難であり、一定の拒否世帯が発生している。

勤労者以外の世帯(無職世帯を除く)については支出のみを調査している⁹⁾。また、毎月の収入・支出に加え、2002年に貯蓄動向調査が統合され、貯蓄・負債も調査されるようになっており、全ての世帯について調査開始後3ヶ月目に調査されている。さらに、世帯員の属性や住居についても詳細な情報が収集されており、社会経済学的な分析を可能としている。

調査は基本的に「家計簿」に記入する方法で調査され、収入の源泉や支出の用途ごとに詳細に分類されて公表されている。家計簿には世帯が自由に記入し、その内容を統計局で集計する際に分類番号を付与して分類する「アフターコード方式」を採用している。この方法では、事前に調査する項目を限定する必要がないため、新製品の登場などによる家計行動の変化を的確に把握することができる。家計調査は、消費者物価指数のウエイト算出に利用されるため、記入項目を事前に限定しない調査方式を採用することは極めて重要である。一方で、自由記述であるため調査世帯の負担は大きく、また記入の必要性に関して誤解の余地があり記入漏れが生じる可能性がある。こうした調査方法の性質により、家計調査のサンプルに偏りがあるという批判や、高額商品の記入漏れが発生していると考えられる原因となっているのである。

家計簿は毎月2回集められるが、公表されるのは月ごとに集計された結果である。ただし、2000年以降は、日別の支出についても公開している。各月の調査結果は、翌月末に速報として公表され、翌々月に「家計調査報告」として刊行されている。また、単身世帯及び総世帯の家計収支に関する結果、貯蓄・負債に関する結果は四半期ごとに公表されている。

3. 家計調査に対する批判

3.1 標本誤差に対する批判

家計調査に関する議論を概観してみると、最大のユーザーであり批判者でもあるのが市場の「エコノミスト」である。その批判は、総合研究開発機構(2008)に集約されており、その中で「市場分析専門家」に対するアンケートとして自由記述された意見として見ることができる。その意味では、家計調査の多様な利用目的のうち、景気動向の指標としての側面に最も多くの関心が集まっていると言える。また、具体的な批判のほとんどは、月次の「変動」が大きすぎることの一点に集中している。例えば、NIRAアンケートの対象となった「市場分析専門家」のうち、記名で家計調査への意見を述べたエコノミスト23名のうち17名が、月次変動の大きさについて批判している。

変動が大きすぎるとは、家計調査の結果が、しばしばGDP統計や毎月勤労統計などの景気指標と説明が困難な乖離をしていることを指す。この批判を、家計調査の標本誤差が大きいという意味で解釈すれば、当然の批判である。総務省統計局の公表によれば、標本調査論

⁹⁾ 1986年以前は、無職世帯についても収入は調査されていなかった。

に基づく消費支出の標準誤差率は2.0%であるのに対し、エコノミストの多くが要求する水準は通常1%ポイント以下であり、より高い精度を要求しているのである。また、その原因については、多くのエコノミストは「サンプルサイズが小さいから」と考え、家計調査の調査設計上の問題であると批判している。

こうした標本誤差に対する批判は、調査世帯を増加させて、より精度の高い統計を作成すればよく、原理的には解決は容易である。しかし、調査の規模の拡大のコストと精度向上の便益の比較、さらには速報性の維持に対する要求も考慮すると、単純に調査世帯を増加させることは最適な対策ではないと考えられる。しかも、会田(2000)で指摘されているように、そもそも毎月の家計の支出は様々な特殊要因により変動するという性質がある。例えば、各月の曜日の構成、天候、公共料金等の引き落としのタイミングなどにより支出は不規則に変動する。こうした変動は必ずしも供給側の統計には表れない要因であり、景気動向とも区別されるべき変動である。つまり、調査世帯を増加させても、説明不能な変動を小さくすることはできない可能性がある。一般に、調査統計においては一定の標本誤差は不可避であり、標本誤差に対応した分析手法の開発がより重要である。

家計調査の調査世帯数を拡大する以外の選択肢として、よりサンプルサイズの大きな家計消費状況調査を活用することが考えられる。家計消費状況調査は、2002年から開始されており、約3万世帯を対象としており家計調査よりも標準誤差率が小さいと考えられる¹⁰⁾。また、「消費や購入頻度が少ない高額商品・サービスなどへの消費の実態を安定的に捉えることを目的」とした統計であり、詳細な項目分類別の支出は利用できないが、家計調査を補完する統計となっている。今後は、消費に関する景気指標としては、家計調査と家計消費状況調査を組み合わせる分析することが望まれる。ただし、現時点では、家計消費状況調査の公表のタイミングは家計調査に比べ遅く、詳細な集計にも対応していない。より安定的な消費動向の指標を構築するためには、家計消費状況調査の公表に関する改善が、最も現実的な対応だと考えられる。

3.2 サンプルセレクションバイアスに対する批判

景気指標としては家計調査の標本誤差が最大の問題であったが、家計行動の分析などの目的のためには2%程度の標本誤差は大きな問題ではない。より深刻な問題は、調査の結果が構造的に「真の値」から異なるという意味での「非標本誤差」である。ここでは、非標本誤差のうち標本である調査世帯の属性が、母集団である日本の家計全体の属性とは異なるという意味での「サンプルセレクションバイアス」について論じる。

前節で見たように、家計調査では標本調査論に基づいた標準的な標本抽出方法を採用している。しかし、実際には多くの世帯が調査拒否をしており、日本の家計部門全体に対して

¹⁰⁾ 回収率が70%程度であるため実際のサンプルは2万世帯程度である。

代表性を持っていないとの批判が多くされている。そうした批判の多くが指摘しているのが、家計調査の調査負担の大きさである。調査世帯は6か月間家計簿を詳細に記入する必要があり、しかも調査開始後の最初の1か月は購入した商品の重さなどの「数量」も報告することになっているため¹¹⁾、他の統計調査と比較しても調査客体に大きな負担となる調査である。そのため、調査に非協力的な世帯や調査に協力することが困難である世帯が調査を拒否したり、調査から脱落する可能性が高いとされている¹²⁾。

これまでに指摘された具体的な偏りとして、1. 公務員世帯が多い、2. 専業主婦世帯が多い、3. 低所得者が少ない、4. 超高額所得者が少ない、が挙げられる¹³⁾。公務員世帯については、他の職業の世帯と比較して調査に協力的であると予想され、実際に調査される世帯での公務員世帯の割合が高くなると考えられている。また、共働き・低所得の世帯は家計簿を継続的に記入することが困難であるため、実際に調査される世帯は、相対的に専業主婦世帯や高所得世帯の割合が高くなると考えられているのである。超高所得者については、プライバシーの観点から、調査拒否を起しやすくと考えられるため、ほとんど調査できていないと言われている。

こうした指摘は、新聞等での「識者の発言」として取り上げられることが多く、論文などのまとまった文章で指摘しているものはほとんどない。そのため、多くの指摘が半ば「都市伝説」の域を出ない状態で繰り返し指摘される状態が続いている。しかし、世帯主の職業や世帯主の配偶者の就業状態などの世帯属性の偏りについては、他の統計と比較することで妥当性の検証は可能であり、ここではそうした議論をまとめることとする。

まず、公務員の割合が高いという疑問については、統計局側から多くの反論が提示されている。その代表的なものが会田(2000)であり、労働力調査と比較することで家計調査の妥当性を論じている。また、ほぼ同様の主張は、現在の統計局のホームページでも解説されている¹⁴⁾。ポイントは、家計調査での「公務員の割合」は世帯主に占める割合であり、人口比率での公務員の割合よりも高くとも必ずしもサンプルの偏りを意味しないということである。ただし、厳密には、概念の調整をしても他の統計よりも1%から2%程度家計調査での公務員世帯の割合が高いことが報告されており「都市伝説」の原因になっていると推測できる¹⁵⁾。とはいえ、公務員世帯の収入・支出の動向は他の勤労者世帯とそれほど異ならなかった

¹¹⁾ 2002年以前は、数量についても6ヶ月報告することになっていた。

¹²⁾ 溝口(1992)では、拒否率が1995年で30.1%、1990年前後で45%程度であるとしている。

¹³⁾ 例えば、日本経済新聞に掲載された記事として、1993年6月12日号「エコノ探偵団：あてになるか？家計調査」、2000年6月19日号「GDP統計改善急げ：問題多い消費と公共投資」、7月14日号「再考・経済統計第7回：消費その2」、2004年1月9日「数字は踊る(1) あいまいを選ぶ—統計の穴、うたかた景気右に左に(けいざい心理学)」、2006年11月10日「最長景気をみる(中) 統計の「中身」手探り(経済予測というもの)」等で指摘されたもの。

¹⁴⁾ 総務省統計局のホームページ>家計調査のページ> Q&A から Q19 を参照のこと(2009年4月1日現在 <http://www.stat.go.jp/data/kakei/qa-1.htm>)。

め、この水準の違いではほとんど影響はないと考えられる。

それに対し、専業主婦世帯が多いという疑問については、容易に検証可能であるにも関わらず批判・反論ともに十分な議論がされていない¹⁶⁾。ここでは、総務省統計局が行なう類似した3つの統計、家計調査・全国消費実態調査・家計消費状況調査を比較することで、この批判の妥当性について検討する。全国消費実態調査は、5年に1度の調査でありサンプルサイズは約6万世帯と大きい。標本抽出や調査の方法は家計調査に類似しているが、調査期間が3カ月と家計調査よりも短い。一方、家計消費状況調査は、調査期間は1年間と長い。特定の品目についてのみの調査であり、事前に調査対象品目が決まっているプリコード方式である。総じて、調査負担は、家計調査>全国消費実態調査>家計消費状況調査の順で重いと考えられている。

それらの3つの統計が利用できる最新の結果である2004年のデータを用いて、世帯属性を比較したものが(表1)である。まず、世帯人員について見ると、3つの統計でほぼ同じ結果となっている。2005年の国勢調査によれば、一世帯あたりの世帯員は2.58となっており、正確に把握できているように見える。しかし、これは、平成12年以降に単身世帯を含めた総世帯の結果を公表するに当たり労働力調査の世帯分布を用いた比例補正した比推定値を公表するようになった結果である¹⁷⁾。すなわち、世帯人員については、そもそも「真の分布」になるように調整されており、一致するのは当然である¹⁸⁾。逆に言えば、以下でみる世帯属性の違いは、世帯人員を調整した後の結果である。

世帯主の年齢については、収入を調査している家計調査と全国消費実態調査では、家計消費状況調査よりも2歳程度低くなっている。世帯主は全て「家計費に充てるための収入を得ている人」と定義されているが、収入を調査していない家計消費状況調査では、例えば同居の親が世帯主になっている可能性がある。家計の行動は年齢と強い相関関係があることが知られており、この点についてはさらなる検討は必要である。

さらに、有業人員についてはより大きな違いが存在している。家計調査の結果が一番小さく、全国消費実態調査、消費状況調査の順になっている。この結果は、二人以上の世帯や二人以上の世帯・勤労者世帯などを見ても同様であり、家計調査において有業人員の多い世帯の割合が過少である可能性が否定できない。例えば、パネルCで世帯主以外の有業者が全て配偶者であるとすれば、就業人員0.2人の違いは、共働き世帯が20%過少になっていることを意味する。これは、公務員世帯のケースより大きな偏りであり、共働き世帯の収入・

¹⁵⁾ 霧島(2000)では、家計調査での公務員世帯の割合が、労働力調査よりも「1.2%も高い」と指摘している。

¹⁶⁾ 日本経済新聞1990年2月21日「経済教室：大きな偏りない家計調査」で、国勢調査と比較して「共働き世帯は...決して低すぎる割合ではない」と論じられている。

¹⁷⁾ 比推定についての詳細については総務省統計局のホームページ>家計調査のページ>「家計調査のしくみと見方」の付録2(<http://www.stat.go.jp/data/kakei/pdf/furoku2.pdf>)を参照のこと

¹⁸⁾ ただし、全国消費実態調査・二人以上の世帯については、世帯人員による比推定をしていない。

表1 3つの世帯調査とサンプルセレクションバイアス(2004年)

パネル A. 総世帯(単身世帯を含む)			
	家計調査	全国消費実態調査	家計消費状況調査
集計世帯数	8,606	58,048	22,530
世帯人員(人)	2.59	2.57	2.60
有業人員(人)	1.19	1.22	1.28
世帯主の年齢(歳)	53.9	53.7	55.5
パネル B. 二人以上の世帯			
	家計調査	全国消費実態調査	家計消費状況調査
集計世帯数	7,742	53,112	20,104
世帯人員(人)	3.19	3.26	3.20
有業人員(人)	1.39	1.49	1.53
世帯主の年齢(歳)	54.1	53.7	56.4
パネル C. 二人以上の世帯・勤労者世帯			
	家計調査	全国消費実態調査	家計消費状況調査
集計世帯数	4,427	31,025	10,180
世帯人員(人)	3.48	3.52	3.49
有業人員(人)	1.63	1.70	1.82
世帯主の年齢(歳)	46.4	46.5	48.7

支出の構造は専業主婦世帯の構造とは大きく異なるため、影響も大きい。さらに、調査負担が大きい順に有業人員が少ないという結果は、調査負担の大きさがサンプルを特定の世帯に偏らせているという批判が妥当している。

この結果から、家計調査の調査負担が大きいと、共働き世帯の調査が困難であり専業主婦世帯の割合が高くなっているという批判には妥当性があると考えられる。この問題に対処するためには、調査方法を改善することや、集計の段階で事後的に専業主婦世帯と共働き世帯の比率を調整する、すなわち比推定の方法を改善する、などの検討が必要だろう。

次に、所得水準に関する偏り、すなわち低所得者や超高所得者が調査から漏れているという批判について検討する。例えば、大竹(2005)が「調査対象になった世帯は、家計簿をつける必要があり、調査を拒否する可能性がある。そのため、機会費用の高い高所得の人や家計簿を付ける余裕のない低所得の人のサンプルが落ちる可能性がある。」と指摘している問題である。

上で見たような世帯属性に関する偏りについては、それぞれ他に信頼できる統計が存在しており、他の統計と比較することが偏りを検証することが容易であった。しかし、低所得者が少ない・超高所得者が少ないなどの「収入」に関する偏りについては、家計調査が主たる情報源であり他の統計と比較することは困難である¹⁹⁾。例えば、賃金構造基本調査や毎月勤労統計などの企業側の統計と比較することは可能であるが、対象となる企業規模に違いがあり、世帯主の賃金に限定したり世帯全体の収入を推定することは困難であることから、必ずしも適切ではない。

家計調査以外の収入に関する情報を持つ家計側の統計としては「国民生活基礎調査」が代表的であり、大竹(2005)は「家計簿をつける必要のない『国民生活基礎調査』の場合は、より低所得や高所得の世帯の回収率が高いと考えられる」と述べている。国民生活基礎調査は3年ごとの調査であり比較可能な年次が限られ、調査方法も家計調査と異なるため直接比較は困難であるが、家計調査の方が低所得者の割合が低いことを指摘している先行研究も存在している²⁰⁾。また、大竹(2005)では国民生活基礎調査で計測されるジニ係数が、家計調査で計測されるジニ係数よりも大きいことを示しており、家計調査のサンプルが所得水準について中央に「偏り」を持っていることが示唆されている。さらに、他の統計との比較ではなく、実際に調査を拒否した世帯の属性を観察することでも評価ができる。例えば、溝口(1992)第4章脚注4は、類似の調査である全国消費実態調査において、年間収入の低い階級で調査拒否世帯の割合が顕著に高いことを指摘している。

こうした結果から、家計調査において低所得者のサンプルが脱落している可能性は否定できない。一方、超高所得者については、そもそも母集団が小さいため偏りを検証することは困難であり、先行研究でもほとんど検討されていない。溝口(1992)は「世帯数の分布からみれば低所得階層の行動が過少にしか反映されていないが、その欠点は金額的に大きい高額所得者が標本に採用されにくいという標本調査の特性によって相殺されてきた」と指摘しており、「平均値の動きが比較的良好に家計の消費行動をフォローしてきたこと」が高所得者が調査世帯からもれていることの証拠と考えている。実際、浜田(2007)はSNAの可処分所得と家計調査の可処分所得を比較して、両者の説明できない相違は5%程度であることを示しており、所得水準に関するサンプルの偏りは「平均」に関する限り大きな問題となっていないことが示されている。その意味で、低所得者と超高所得者の両者がサンプルから脱落しているという議論には説得力がある。

¹⁹⁾ 家計調査の標本抽出の単位が住居であることから、住居を持たない世帯、いわゆるホームレス、などについては調査対象から漏れており、そうした人々は一般に低所得者であるため、その意味においては低所得者の割合が低いことは否定できない。しかし、こうした世帯を家計調査の枠組みの中で調査することは現実的ではないため、ここでは考慮しない。

²⁰⁾ この点について、溝口(1992)では、山田(1990)が示しているとの記述があったが、原論文では確認していない。

繰り返しになるが、この所得に関する2つの偏りは相殺されるため、平均には大きな影響を与えない。しかし、近年の格差論争に顕著なように、世帯の経済水準については平均だけではなく分布そのものが関心の対象となっており、調査世帯の代表性を高める努力は必要である。世帯ベースの所得分布に関して、家計調査比較可能な月次の統計である家計消費状況調査が利用可能となったことで、今後の分析の進展が期待できる。ただし、現状において、格差についての分析や、低所得者層および超高所得者層についての分析には、他の統計も併せて活用することが適切である²¹⁾。

3.3 項目別の測定誤差に対する批判

前節で見たように、家計調査のサンプルに偏りがあるという批判があり、その一部には妥当性があることが分かった。それに対し、ここでは、問題があると指摘されている家計調査の調査項目に注目して、その妥当性を確認する。もちろん、偏ったサンプルを対象に調査をすれば各項目について偏った結果となるが、ここではサンプルによる偏りとは別に、固有の原因により偏りがあると考えられている項目、具体的には、高額消費全般・財産所得・貯蓄率について論じる。

家計調査の支出側の項目のうち、単価の大きい「高額消費」は、毎月の購入頻度が少なく結果が安定しないということが知られている。高額消費とはリフォームや増改築などの住宅設備関係・家具・自動車・冠婚葬祭費などであり、購入頻度は極めて小さく9,000世帯程度の調査では月次の誤差が大きいことが予想できる。すなわち、これら的高額消費をした世帯が調査対象になるかどうかによって支出額が大きく変動してしまい、解釈困難な変動を生み出されるのである。しかし、購入頻度が低いことは「標本誤差」を生み出す原因にはなるが、それだけでは結果の偏りを生み出す「非標本誤差」の原因とはならない。その意味では、先行研究の多くが家計調査の高額消費が過少であると考えていた根拠は明らかではない。

高額消費の中でも、典型例として自動車購入がしばしば取り上げられる²²⁾。家計調査における自動車購入に偏りがあるとする根拠として、新車販売(登録車)台数との比較がされている。梅田・宇都宮(2003)では「自動車については、高額で購入頻度が低く家計調査のデータを用いることができないことから、家計消費支出の動向を把握する上では必ず用いられる指標である」とされている。しかし、家計調査の「自動車購入費」と新規登録・届け出台数を比較するのは多くの問題がある。その理由は、まず、第1に登録時期と車の受け取り・支払時期のズレである。すなわち、家計調査が完全に自動車の取引を把握していたとしても若干のタイムラグが生じる。第2に、新車登録の中になんかの数の営業車・社用車が含まれているからである。法人企業による営業車・社用車は家計調査の対象外であり、供給側

²¹⁾ 例えば、超高所得者層については橘木・森(2005)、低所得者層については橘木・浦川(2006)を参照。

²²⁾ 日本経済新聞1999年7月7日号「観潮台：消費動向の盲点」でも指摘されており、NIRA(2008)のアンケートでも数名が指摘している。

の統計との差を生み出す。第3に、重要であるにも関わらずよく知られていない事実として、家計調査では中古車の下取り価格が相殺されていることである。概念的には、中古車の下取りによる収入は資産の売却として計上すべきであるが、実際の取引においては、新たに購入する自動車の実質的な値引きとして運用されることも多い。そのため、家計調査では、新車の購入代金から中古車の下取り代金を差引いた金額を「自動車購入費」として計上しているのである。この取扱いが適切であるかは議論が必要であるが、単純な供給側の統計との比較はできないのである。

他の「高額消費」に関しては、誤差が大きいという批判はあるが、その原因についてはほとんど指摘されておらず、偏りがあるかどうかについても検討されてこなかった。それに対し、宇南山(2009)では、「家計消費状況調査」を用いて偏りに関して検証をしている。

家計消費状況調査は、上でも述べたように「購入頻度が少ない高額商品・サービスなどへの消費の実態を安定的に捉えることを目的」とした統計であり、対象となる財サービスを事前に調査票に示し消費金額のみが書き込まれる「プリコード方式」で調査されている。調査品目は、基本的に単価が高く購入頻度の低い財・サービスであり、自動車やパソコン・デジタルカメラ・冷蔵庫・エアコン・家賃・学校の授業料・旅行・冠婚葬祭費などが該当する。すなわち、家計消費状況調査は、家計調査に比べサンプルサイズが大きく、従来から家計調査の問題点として指摘されていた品目に特化した統計であり、少なくとも調査対象品目に関してはより正確な支出金額を把握していると考えられる。

宇南山(2009)では、家計消費状況調査の調査対象品目のうち家計調査と比較可能な品目に対する支出額の合計を比較しており、その結果をまとめたものが(表2)である。この表では、2002年から2006年までの年平均の支出額を2つの統計で計算し、家計調査の支出額を家計消費状況調査の支出額で除した「比率」が計算されている。この比率は、標本誤差を除けば100%となるはずで、家計調査の結果が「過少」であるならば100%を下回ることになる。実際に計算された比率は、2002年から2006年まで約70%前後で安定しており、家計調査の支出額が家計消費状況調査の支出額よりも低い水準であることが分かる。さらに、(表3)は、比較可能なすべての品目に対して、対応関係を示しこの比率を計算して、5年分を平均したものである。この表から、固定電話通信料・インターネット接続料・家賃など定期的に支払う通常の支出はほぼ両統計で一致しているのに対し、ステレオセット・食器戸棚・挙式披露宴費用・葬儀法事費用など、購入頻度が低く購入単価の高い財・サービスでは、家計調査の支出額が家計消費状況調査の支出額を大幅に下回っていることが分かる。すなわち、家計調査で把握されている高額消費への支出額は、誤差が大きいだけでなく実際の支出額を過少評価する偏りを持っているのである。

ただし、ここで示された高額消費の過少性がなぜ発生するかについては検討の必要がある。家計消費状況調査との比較に加え、もう一つの世帯調査である、全国消費実態調査を用

表 2 家計調査と家計消費状況調査の比較

	調査年					
	2002	2003	2004	2005	2006	2007
家計調査 (1)	101,824	100,562	107,441	105,235	103,492	104,864
家計消費状況調査 (2)	149,428	147,414	148,996	150,474	152,851	150,772
比率 ((1)/(2))	68%	68%	72%	70%	68%	70%

(1 世帯あたり・月平均・円)

いて同様に比率を計算すると、その比率は 94%ととなり 2つの調査の結果は大きく異なることが分かる。全国消費実態調査のサンプルサイズは家計消費状況調査よりも大きい約 6 万世帯であることを考慮すると、高額消費が過少であることはサンプル数が少ないことが原因ではなく、むしろ家計調査と全国消費実態調査では共通であるが、家計消費状況調査では異なる性質が原因と考えられる。一つには、上で述べた家計調査のサンプルセレクションバイアスの影響が考えられる。例えば、共働き世帯では耐久消費財の購入頻度が高いとすれば、家計調査と家計消費状況調査の違いは、調査世帯の世帯属性の違いで説明できるはずである。また、調査方法が原因となる可能性も考えられる。すなわち、家計消費状況調査で採用されているプリコード方式の方が記入漏れが少ないという可能性である。

しかし、どちらにしても、本質的には家計調査で家計簿を自由記入する方式を採用していることが原因と言える。自由記入の記入負担がサンプルセレクションを引き起こし、さらに記入漏れも発生させていると考えられるからである。ただし、家計調査が消費者物価指数のウェイト算出に使われることを考慮すると、家計調査の調査方法をプリコード方式に切り替えることは困難である。さらに、プリコード方式では、もともと意図していない項目が含まれる危険性もある。例えば、自動車購入に関して、税や手続き費用を含んだ総額で回答されてしまう可能性があり、項目としては過大になる可能性がある。その意味でも、アフターコード方式で、より正確に高額商品への支出を把握する方法については検討が必要である。

つぎに、財産収入について考える。財産収入とは家計の資産から生み出される所得であり、重要な所得の源泉である。家計調査においては、資産からの所得は金融資産・実物資産の区別はなく「財産収入」としてまとめて表象されているが、家計調査年報「収支項目分類の例示」によれば、「金融資産、土地及び無形資産（著作権・特許権など）の賃貸によって発生する収入、預貯金利子 貸金利子 株式配当金 公社債利子 立看板（広告）の貸地料 土地の権利金収入（家計調査年報・家計収支編・平成 19 年 p.481）」によって構成されている。

この「財産収入」が家計調査においては正確に把握されていないという批判が存在して

表3 品目別の家計調査と家計消費状況調査の支出金額の比

家計消費状況調査の品目名	家計調査の品目名	比率
携帯電話(携帯電話・PHS)使用料	該当あり	73%
固定電話使用料	該当あり	100%
インターネット接続料(プロバイダー料)	インターネット接続料	92%
同上(プロバイダー料と通信料)		
ケーブルテレビ受信料(受信料)	ケーブルテレビ受信料	135%
同上(受信料とインターネット接続料)		
衛星デジタル放送受信料	該当無し	
たんす	該当あり	46%
ベッド	該当あり	40%
布団	該当あり	49%
机いす(事務用学習用)	該当無し	
食器戸棚	該当あり	35%
応接セット	該当あり	41%
ピアノ	該当あり	26%
背広服	該当あり	51%
婦人用スーツワンピース	該当あり	44%
和服(男子用婦人用)	該当あり	39%
自動車(新車)	自動車購入	31%
自動車(中古車)		
自動車保険料(自賠責)	該当あり	51%
自動車保険料(任意)	該当あり	81%
自動車以外の原動機付輸送機器	自動車以外の輸送機器購入	29%
自動車整備費	該当あり	30%
家屋に関する設備費工事費修理費(内装)	設備修繕維持	74%
家屋に関する設備費工事費修理費(外装)		
給排水関係工事費	該当あり	28%
庭・植木の手入れ代	該当あり	41%

品目別の家計調査と家計消費状況調査の支出金額の比 (続き)

家計消費状況調査の品目名	家計調査の品目名	比率
家賃	民営家賃+公営家賃+給与住宅家賃	110%
宅地の地代	地代	55%
冷蔵庫	該当あり	58%
洗濯機	該当あり	59%
エアコンディショナ	該当あり	55%
ミシン	該当あり	39%
ステレオセット	該当あり	35%
パソコン	該当あり	63%
パソコン用周辺機器ソフト	該当無し	
移動電話機 (携帯電話機 PHS の本体価格と加入料)	移動電話	53%
インターネット接続機能付固定電話機	該当無し	
ファクシミリ付固定電話機	該当無し	
携帯情報端末 (P D A)	該当無し	
デジタル放送チューナー内蔵テレビ	テレビ	86%
デジタル放送チューナー内蔵テレビ以外のテレビ		
デジタル放送用チューナーアンテナ	該当無し	
インターネット接続機能付テレビゲーム機	該当無し	
デジタルカメラ	カメラ	44%
デジタルカメラ以外のカメラ		
デジタルビデオカメラ	ビデオカメラ	44%
デジタルビデオカメラ以外のビデオカメラ		
インターネット接続機能付カーナビゲーション	該当無し	
歯科診療代	該当あり	63%
歯科以外の診療代	医科診療代	70%
出産入院料	該当あり	52%
出産以外の入院料	該当あり	47%

品目別の家計調査と家計消費状況調査の支出金額の比 (続き)

家計消費状況調査の品目名	家計調査の品目名	比率
国公立授業料等 (幼稚園～大学, 専修学校)	授業料等	76%
私立授業料等 (幼稚園～大学, 専修学校)		
補習教育費	該当あり	80%
自動車教習料	該当あり	39%
航空運賃	該当あり	54%
宿泊料	該当あり	80%
パック旅行費	該当あり	72%
国内パック旅行費	該当あり	82%
外国パック旅行費	該当あり	51%
挙式披露宴費用	婚礼関係費	24%
葬儀法事費用	葬儀関係費	40%
信仰関係費	信仰・祭祀費	52%
仕送り金	該当あり	115%
贈与金	該当あり	104%
合計		69%
支出総額	消費支出	84%

いる。例えば、岩本・尾崎・前川(1996) および浜田(2007) では家計調査の1世帯あたりの金額に世帯数を乗じて「マクロ」の財産収入を計算しSNAの該当項目と比較しているが、家計調査から計算された数値はSNAの数値の約5%に過ぎず、他の収入項目と比較しても家計調査とSNAの乖離が特に大きいことを指摘している²³⁾。また、高山他(1989)では全国消費実態調査の資産・負債編を活用することで「保有する金融資産に比べて利子・配当所得等が過少に記載されている」可能性を指摘している。

実際に、SNAの財産所得を1世帯あたりに換算して家計調査の結果と比較してみると、家計調査の結果が過少であることは明らかである。SNAでは1990年の財産収入は約40兆円であり国勢調査での世帯数は4104万世帯となっている。すなわち、1世帯あたりの年間97万円、月平均で約8.1万円の財産収入が記録されているのに対し、家計調査の「二人以上世

²³⁾ 世帯調査における財産所得の過少性については、海外でも指摘されている。例えば、Atkinson *et al* (1995) および Banks and Johnson (1998) を参照。

帯(勤労者世帯と無職世帯の合計)」で計算される平均の財産収入は月0.3万円である。同様に、2005年では、SNAでの1世帯あたり財産収入が1.3万円であるのに対し、家計調査はわずかに991円となっている。2つの統計には、概念上の違い存在しており、厳密な比較は困難であるが、バブルが崩壊してから10年が経過した2000年以後も一貫して家計調査の財産収入が低い水準にあり何らかの調査実務上の問題が存在すると考えられる。

岩本・尾崎・前川(1996)では、財産収入(典型的には利子所得)が再投資された場合(例えば定期預金が自動延長された場合)などでは、財産収入の記入もれが生じやすいと論じている。財産収入の記入もれは、高額商品のケースなどと異なり、家計自身が認識していないケースも多いと考えられるため、調査方法の改善は困難であると考えられる。しかし、今後、高齢化が進展することで、家計の収入源が労働所得から資産所得に移行していくことが予想されたため、資産所得を正確に把握することの重要性は高まるため、何らかの対応は必要であろう。

最後に、家計調査の貯蓄率に対する批判について述べる。日本においては、SNAで計算される家計貯蓄率と家計調査から計算される家計貯蓄率との間に大きな乖離があることが知られている²⁴⁾。(図1)はSNAと家計調査の貯蓄率の推移を示したものであるが、この図から直近では両統計の貯蓄率が20%以上も乖離していることが分かる。特に、1980年以降は、貯蓄率の水準のみならず変化の方向も異なっている。

この問題は、家計調査に対する批判の中で、学術的には最も注目された問題であり、その乖離の原因を解明するために多くの研究が行われてきた(例えば、植田・大野, 1993; 村岸, 1993; 岩本・尾崎・前川, 1995・1996; 中村, 1999)。しかし、これまでのところ乖離の原因は完全には解明されておらず、家計調査の信頼性に対する一つの大きな問題となっている²⁵⁾。

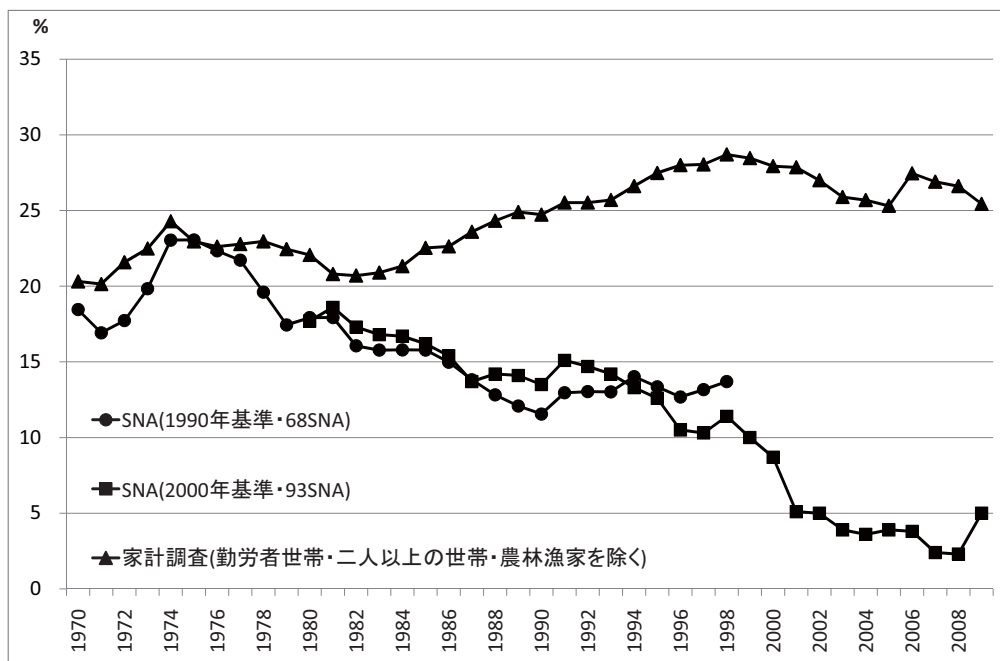
多くの先行研究では、2つの統計の対象範囲の違いや貯蓄の概念的な違いにより、貯蓄率の乖離を説明しようとしてきた。もっとも包括的に論点を整理した岩本・尾崎・前川(1995・1996)では、おおむね貯蓄率の乖離の約70%はそれぞれの統計の制度的・概念的な違いで説明できることを示した。しかし、残された乖離の30%の部分についてその原因は明らかではないと結論付けていた。その後、梅田・宇都宮(2003)でも「1990年代半ばに議論となった際には、標本調査である家計調査の誤差を示唆する結果が相次いだが、コンセンサスは得られていない」としている。

それに対し、宇南山(2009)では、先行研究ですでに指摘されていた要因に加え、上で述べた高額商品への支出・財産所得が過少であることを乖離の原因であると特定した。さらに、

²⁴⁾ Demery and Duck (2006) は、英国について、マクロ統計とミクロ統計から計算される貯蓄率を比較し、長期的には大きな違いがないことを確認している。

²⁵⁾ 1990年代の貯蓄率の乖離に関する論争をまとめたものとして谷沢(1999)が便利である。

図1 SNAと家計調査の貯蓄率の推移



家計消費状況調査や貯蓄動向調査などの統計を活用することで、これらの項目の過少性を補正した上で貯蓄率を比較すると、2つの統計の乖離はほぼ解消されることを示した。言い換えれば、よく知られている貯蓄率の乖離の問題は、その他の家計調査の問題点を反映した副次的な問題と考えられるのである。

以上、家計調査の結果に問題があると批判される項目として、高額商品・財産所得・貯蓄率について、批判の妥当性について検討した。その結果、実際に、これらの項目では、家計調査の結果が過少であることが明らかになった。しかし、家計調査が複数の目的に利用されていることを考慮すると、調査方法をすぐに変更することは困難であり、実際に家計消費状況調査など補完的な統計を調査することで対応が進められている。その意味で、家計調査の結果を利用する際に、特にこれらの項目について分析する場合には、他の統計の結果で補完しながら利用することが重要である。

3.4 その他の改善について

その他の家計調査に関する批判のうち、最も重要なものとして、個人営業世帯等の収入を調査していないことが挙げられる。家計調査における収入の動向や貯蓄率を他の統計と比較する際に、個人営業世帯等が利用できないことは大きな制約となっている。現在、個人営業世帯の収入を調査しないのは「月々の収入を営業上の収入と家計収入に切り離してとら

えることが難しいため」とされている²⁶⁾。しかし、こうした個人営業世帯についても年間収入については調査しており、調査期間が2週間である英国でも個人営業世帯の所得を調査していることから、調査そのものは可能であると考えられる²⁷⁾。今後、個人営業世帯の収入も調査対象とすることは重要な課題であろう。

家計の代表性について関連する点として外国人世帯の問題についても対応が必要である。現在、家計調査の調査対象に含まれていない世帯は(1) 料理飲食店、旅館又は下宿屋(寄宿舎を含む。)を営む併用住宅の世帯；(2) 賄い付きの同居人がいる世帯；(3) 住み込みの営業上の使用人が4人以上いる世帯；(4) 世帯主が長期間(3か月以上)不在の世帯；(5) 外国人世帯；(6) その他(単身の学生世帯など)であり、日本人の収入・支出の動向を捉えるには十分な範囲である。ただし、今後、外国人世帯が増加することが予想され、日本の家計部門全体を把握するために外国人世帯を調査対象とする必要があると考えられる。しかし、調査員の言語の問題など多くの課題があり、調査方法を含めた早めの対応が必要である。

さらに、よりマイナーな点について、2点述べる。まず第1点は、家計調査において無職の世帯員についてその理由を聞いていない。無職であるという就業状態であっても、通学・専業主婦・引退などの「非労働力」であるか、「失業状態」であるかによって経済学的には大きな違いが存在する。英国や米国などの調査では利用可能な変数であり、高齢化が今後ますます進展すると予想されるため、特に高齢者の失業・引退の意思決定の分析が重要となるため、この区別を導入することは大きな意義があるだろう。

もう1点は、世帯員の学歴に関する情報である。この変数も、他の国では一般的に調査されている項目であるが、日本の家計調査では調査されていない。日本においては、学歴を調査することへの抵抗が大きいと予想できるが、近年の研究では所得や就業さらには支出に関しても学歴の影響が無視できないことを示す研究が蓄積されつつあり、経済社会の重要な変数として学歴を調査するニーズは高まっていると考えられる。

3.5 公表に関する改善について

ここまで、家計調査の収集するデータの品質についての批判とその妥当性を検討してきた。しかし、家計調査において収集されるデータの品質が向上しても、それが利用しやすい形態で提供されないと、利用者にとっては意味がない。特に、経済学の発展や社会構造の変化に伴い、研究者が考える最適な表象が変化してきており、現在公表されている集計方法が家計行動を把握する上で最適な表象となっていない可能性がある。

それに対し、統計局も結果の公表についての改善もしばしば行なっている。例えば、1980

²⁶⁾ 総務省統計局「家計調査のしくみと見方」p.7より(<http://www.stat.go.jp/data/kakei/pdf/mikata1.pdf>を参照)。

²⁷⁾ Banks and Johnson (1998) では、個人営業世帯の収入の信頼性は低いと述べている。

年時点では5%程度であった無職世帯の割合が、その後の失業率の上昇や高齢化の進展による引退者の増加などによって2005年時点では25%をこえるまでに変化している。その変化に対応して、「無職世帯」に関する詳細な集計結果が、統計局のホームページ上で公開されるようになってきている。また、既存の集計をより容易に可能とするようなデータベース機能も利用可能となっている。さらに、2009年4月から統計法が改正され、データの2次利用の一環として、オーダーメイド集計と呼ばれる研究者の希望に応じた特別集計が、これまでの手続きよりも格段に容易に可能になる可能性もある。しかし、一般的な統計の利用者が入手する公表データを適切に集計することは依然として重要である。

現在、家計調査の家計収支編では、収入階級別・世帯人員別・世帯主の年齢階級別・世帯主の職業別・住居の所有関係別の集計が公表されている。さらに、勤労者世帯については世帯主の産業別・勤め先企業規模別のデータについても公表されている。特に、収入階級別の結果は五分位・十分位が別に用意されるなど、詳細な分析が可能となっている。しかし、この集計パターンは、基本的に過去50年の間変化しておらず、経済環境の変化や経済学の進展に対応していない。

旧来型のケインズ型の消費関数に基づけば、消費の最大の決定要因は同時点での所得である。実際、家計調査でも所得水準別の集計が最も重視されている。しかし、近年の経済学では、所得・消費などの経済活動のライフサイクルを通じた変化に注目するようになってきている²⁸⁾。そのため、現在では、年齢別の集計に対する関心が高まっている。

さらに、消費のライフサイクル仮説に基づけば、消費は家計の期待生涯所得を反映しており、1時点の所得や資産よりも正確な経済厚生 の尺度となると考えられる。その意味で、経済的な豊かさの水準で家計を分類するのにより適した分類は「消費水準別」の集計であると考えられており、大竹(2005)でも「年齢・消費階級別データの整備が望まれる(p.86)」と指摘されている。このように、年齢別・所得水準別のクロス集計や年齢別・消費水準別のクロス集計を整備することへの潜在的な需要は大きく、集計の方法も再検討する必要がある。

また、年齢別の集計をする際に「1歳刻みの年齢別」の集計をすることも検討の価値がある。もちろん、経済分析をする際には、現行の5歳刻みで十分であるケースがほとんどであるが、1歳刻みのデータには「生年別コーホート」データを構築できるという利点がある。繰り返しクロスセクションデータが利用できる場合には、同一生年の集団を追跡することで、パネルデータに近い分析が可能になることがDeaton(1985)で提示されており「擬似パネルデータ(Pseudo Panel)」と呼ばれている。多くの海外での先行研究では、サンプルサイズの関係で、生年を3から5年ごとにプールしていたが、日本の家計調査を用いることで1歳刻みの擬似パネルデータを構築することができる。1歳刻みのデータが利用できれ

²⁸⁾ 消費の決定理論の発展については、例えば、宇南山(2006)を参照。

ば、例えば政策的な制度の変更など、年齢が決定的に重要な経済現象を分析する際に貴重な情報を提供することができる。

さらに、サンプルセレクションバイアスの節で述べたが、家計調査の調査世帯に偏りがあるとの批判は常に存在している。そのため、どのような世帯の拒否率や脱落率が高いのかを検証することは重要である。一般に、サンプルセレクションバイアスの評価には「調査できなかった世帯」と比較が必要なため検証は困難であるが、家計調査においては「準調査世帯票」を用いることで、一定の分析は可能である。

家計調査において、最初に選択された世帯の調査への協力が得られない場合には、その世帯の基本的な特性を「準調査世帯票」によって把握している²⁹⁾。しかし、この準調査世帯票に関する情報は、ほとんど公開されていない³⁰⁾。家計調査は、統計法上の基幹統計であり調査拒否は許されておらず、調査拒否に関する情報を公開することは調査に協力する世帯への心理的影響があるため、慎重に公開する必要がある。しかし、家計調査を適切に利用するためには不可欠な情報であり、より容易にアクセスできるように改善する必要があるだろう。例えば、類似の統計であり基幹統計でもある全国消費実態調査に関しては、準調査世帯票の概要についても公開されており、同様の対応は可能と考える。

このように、家計調査は、調査段階だけではなく、集計・公表段階でも様々な改善が可能である。今後、新統計法の下でマイクロデータそのものが活用される機会が増えると考えられるが、一般的な統計利用者のためにも、理論的な発展を反映した形式で集計・公表されることを期待する。

4. 家計に対する調査全般に関する展望

ここまで見たように、日本の「家計調査」は様々な問題を抱えていることは明らかである。しかし、家計行動の繰り返しクロスセクションデータとして評価するならば、質・量ともに世界最高水準にあると考えられる³¹⁾。例えば、英国の FES は、年間のサンプル数は約 10,000 であり、日本の約 10 分の 1 である。これは、日本と英国の人口比率を考慮しても大きな違いである。調査項目はほぼ同等であるが、FES の調査期間は 2 週間であり、1 ヶ月単位で 6 ヶ月間同一世帯を調査している日本の家計調査の方が優れたデータである。また、米国の家計調査に相当する統計である CES は、調査が開始されたのは 1980 年からと遅れており、1 年間のサンプル数は 7,000 程度である。CES では同一世帯を四半期ごとに 1 年間調査しているが、日々の支出を家計簿をつけて調査するのは 1 週間だけであり、それ以外は耐

²⁹⁾ 準調査世帯票の具体的な内容については <http://www.stat.go.jp/data/kakei/form5.pdf> を参照。

³⁰⁾ 総務省統計局にある統計図書館で閲覧による公表の措置がとられている。この点については、総務省統計局の大貴裕二氏の指摘を受けた。記して感謝したい。

³¹⁾ 世界での「家計調査」に対応する統計の利用可能性については、重川 (2007) が包括的に調査している。

久財等に限定されている。その意味では、家計調査は、日本の貴重な統計であり、改善をしながら維持していくことが重要と考えられる。

しかし、より広く「世帯を対象とした家計行動に関する統計」の利用可能性について考えると、日本の現状が必ずしも優れているとはいえない。特に、近年の経済学・計量経済学の手法の発展に伴い、家計調査のような家計行動の繰り返しクロスセクションだけでは分析できない側面が明らかにされ、同一の家計を長期間追跡したデータとしてのパネルデータに多くの関心が集まっている。国際的に見ても、まだ実験的な段階にあるが、すでに多くのパネルデータが利用可能となってきている。現在、利用可能なパネルデータのうち最も代表的なのが、米国の Panel Study of Income Dynamics(PSID) である。PSID は、1968年から調査が開始され、断続的にサンプルが追加されているが、基本的に同一家計を追跡して所得や食料費に関するデータを収集している。さらに、英国の British Household Panel Survey が1991年から調査開始されており、それを活用した分析も公刊され始めている。また各世帯を8四半期合計2年のみではあるがスペインの Encuesta Permanente de Consumo(ECP) なども調査されている³²⁾。日本の「家計調査」をパネルデータとして見ると、各月のデータを世帯ごとにマッチングすることで家計ごとに6ヶ月追跡が可能であり、短期のパネルデータを作成することができる。しかし、長期のパネルデータがすでに利用可能となっている米国・英国の現状と比較すると、大幅に遅れているといわざるを得ない。

しかも、現在の家計調査を短期パネルデータとして利用する際に、大きな問題が存在していることが知られている(宇南山, 2008)。それは、家計調査の世帯ごとのコード付けが、パネルデータとしての利用を前提としていないため、同一世帯の各月のデータをマッチングすることが必ずしも容易でないことである。家計調査をパネルデータとして利用した研究のさきがけである林(1986)によって、一定の方法が提示され、その後の研究もその方法(もしくはその派生的な方法)を踏襲している(Kohara, Ohtake, and Saito, 2002; 堀・清水谷, 2005; Hori and Shimizutani, 2005a, 2005b, 2006; Shimizutani, 2006; 美添・荒木, 2006)が、宇南山(2008)はそれらの論文で用いられた方法では正しいパネルデータが構築できないことを指摘している。さらに1980年から2005年のデータについてはその解決方法も示しているが、依然として、コード付けの方法は改善されておらず、パネルデータとしての利用を制約している。少なくとも、現状で利用可能な短期のパネルデータを活用するという観点から、世帯のコード付けの方法を改善する必要があるだろう。

さらに、本格的なパネル調査が家計調査の拡張のような位置づけで可能であれば、大きな

³²⁾ 退職前後の高齢者を対象としたパネルデータとして、米国では Health and Retirement Study (HRS)、イギリスでは English Longitudinal Study of Ageing (ELSA)、大陸ヨーロッパでは Survey of Health, Ageing and Retirement in Europe (SHARE) が調査されている。ただし、これらに対応する統計としては、日本版 HRS が市村英彦東京大学教授および経済産業研究所を中心としたグループによって調査が開始されている。

改善となるだろう。日本にも、長期のパネルデータである財団法人家計経済研究所の「消費生活に関するパネル調査」が存在しているが、日本全体を対象とした標本設計となっていないため、例えば政策分析には利用が困難である。また、家計調査よりは長期のパネルデータという意味では、2000年まで家計調査の付帯調査として実施されていた貯蓄動向調査が存在していた。貯蓄動向調査は、家計調査の調査世帯の一部が調査対象であったが、1年後に同一世帯を再調査するという設計になっていた。そのため、貯蓄・負債に関しては2年分のパネルデータが利用可能となっていたのである。この性質が家計調査に統合される際に廃止されてしまったことは、極めて残念である。いずれにしても、パネルデータは、データが蓄積してこそ価値があるため、一刻も早い対応が重要である。

パネルデータを構築することと関連して、家計調査の調査方法の検討も必要である。パネルデータでは、同一世帯を調査することが重要となるため、調査負担を大きくすることは避けなければならない。また、多くの家計調査の課題も調査負担に関連していたことを考慮すれば、調査負担の軽減は現行の家計調査の改善にも資するだろう。すでに繰り返して述べたが、家計調査の目的からすると、アフターコード方式を維持することは極めて重要であり、プリコード方式に変更することによる負担の軽減は困難である。一方で、米国のCESと同様に記憶による調査をすることにも問題がある。溝口(1992)によれば、総理府統計局の「家計調査調査方法研究会」で、家計簿方式から記憶による調査に切り替えて負担を軽減することが検討されていたが、結果として、記憶方式の誤差が大きかったため導入されなかったとされており、慎重な検討が必要である。例えば、民間の調査の一部で活用されているPDAなどの機器を使った負担軽減などの導入は検討の価値があるだろう。

より長期的な課題としては、他の統計との照合可能性の確保についての検討も必要と考える。米国では、原則的に、世帯は社会保障番号などで管理されている。そのため、例えば、税務関連の情報との照合可能性が確保されており、潜在的には生涯所得のプロファイルなどを知ることができる。もちろん、日本において即座に税務統計などを照合することは、法律的にも調査客体の心理的負担の面でも問題が多いが、十分な議論の上で将来的には検討の価値がある。そのためにも、他の統計との照合可能性を確保しておくことは重要な論点となるだろう。

この節で指摘した点は、多くがすぐに対応できる問題ではないが、研究者と統計当局の間で議論を深め実現する価値があると考えられる。統計当局には、多くの短期的な課題の解決とともに、より長期的な課題への積極的な取り組みを期待したい。

5. 結論

本稿では、家計調査に対する批判を概観し、その妥当性を検討した。景気動向に注目しているエコノミストの家計調査への評価は極めて低い。その理由は、家計調査が数%程度の標

本誤差を想定しているのに対し、景気判断においてははるかに高い精度を要求されているためであった。しかし、調査の規模を拡大することで解決することは、コストの面で困難であると考えられ、その有効性も明らかではない。規模の大きな家計消費状況調査が開始されており、高い精度が必要とされる分析には、そういった他の統計で家計調査を補完することで対応する必要がある。

いわゆるエコノミストが標本誤差を問題として家計調査を批判していたのに対し、学術的な関心を持つ研究者は構造的な偏りを持つという意味での非標本誤差について指摘していた。具体的には、サンプルセレクションバイアスと調査項目ごとの測定誤差であり、本稿ではこれらを中心に検討した。

サンプルセレクションバイアスについては、公務員世帯の割合が過大である、共働き世帯の割合が過少である、低所得者・超高所得者の割合が過少であると指摘されていた。公務員世帯が多いという批判は妥当ではなかったが、専業主婦世帯の割合が過大であるという指摘は一定の妥当性を持っていた。所得水準ごとの割合については検証が困難であるが、偏りがあることを示唆する結果が提示されていた。

一方、測定誤差については、耐久消費財などの高額商品や財産所得に記入漏れが発生していると指摘されており、また貯蓄率の動向が他の統計と整合的でないと指摘されていた。高額商品については、2002年から調査が開始された家計消費状況調査を用いて検証すると、家計調査の結果が過少であった。また、財産所得についても多くの記入漏れが発生していることが示唆された。さらに、貯蓄率については、先行研究で指摘されていた要因に加え、ここでも検証した高額商品や財産所得の過少性が原因で偏った結果になっていたことが分かった。

家計調査は、繰り返しクロスセクション統計としてみれば、国際的にも最高水準の統計であるといえる。しかし、経済学の実証に伴う経済学者の関心の対象の変化には十分には対応できていない。特に、家計の動学的な行動を描写するには、同一の個人の行動を追跡したパネルデータの重要性が増してきている。また、近年のプライバシー意識の高まりを受け、個人情報情報を外部に出したくない世帯が増えていることを考慮して何らかの対応が必要であろう。そのために、今後は、現行の家計調査の改善とともに、世帯の収入・支出を世帯調査からどのように把握していくかについての体系的な検討が必要であると考えられる。

謝辞

本稿の作成に当たり、総務省統計局の大貫裕二氏、佐藤朋彦氏には多くの情報を提供していただいた。記して感謝したい。

参考文献

- 会田雅人 (2000) 「家計調査批判への反論」『週刊東洋経済』8月12-19日号
Atkinson, A. B., L. Rainwater, and T. M. Smeeding (1995) *Income Distribution in OECD Countries: Evidence*

- from the Luxembourg Income Study, Paris, OECD.
- Banks, J. and P. Johnson (1998) *How Reliable Is the Expenditure Survey?: Trends in Incomes and Expenditures over Time*, Institute of Fiscal Studies: London.
- Deaton, Angus (1985). "Panel Data from Time Series of Cross-Sections," *Journal of Econometrics* vol. 30(1-2), pp. 109-126.
- Demery, D. and N. W. Duck (2006) "Demographic Change and the UK Savings Rate," *Applied Economics*, vol 38, pp. 119-136.
- 浜田浩児 (2001) 『93SNA の基礎: 国民経済計算の新体系』 東洋経済新報社
- 浜田浩児 (2007) 「家計調査とマクロ統計の比較と整合化」 御船美智子/家計経済研究所編 『家計研究へのアプローチ』 ミネルヴァ書房 第 4 章
- 林文夫 (1986) 「恒常所得仮説の拡張とその検証」 『経済分析』 vol.101 pp.1-23.
- Hori, M. and S. Shimizutani (2005a) "Price Expectations and Consumption under Deflation: Evidence from Japanese Household Survey Data," *International Economics and Economic Policy*, vol.2, pp.127-151.
- Hori, M. and S. Shimizutani (2005b) "Consumer Response to the 1994 Tax Cut: Evaluating the Japanese First Tax Cut in the 1990s," *Hitotsubashi Journal of Economics*, vol.46, pp.85-97.
- Hori, M. and S. Shimizutani (2006) "Did Japanese Consumers Become More Prudent During 1998-1999? Evidence From Household Level Data," *International Economic Journal*, vol.20, pp.197-209.
- 岩本康志・尾 哲・前川裕貴 (1995) 「『家計調査』と『国民経済計算』における家計貯蓄率動向の乖離について (1): 概念の相違と標本の偏りの問題の検討」 『フィナンシャルレビュー (財務省財政金融研究所)』 第 35 号
- 岩本康志・尾 哲・前川裕貴 (1996) 「『家計調査』と『国民経済計算』における家計貯蓄率動向の乖離について (2): ミクロデータとマクロデータの整合性」 『フィナンシャルレビュー (財務省財政金融研究所)』 第 37 号
- 霧島和孝 (2000) 「GDP 統計論争 (2): GDP 統計 私の抜本改革案」 『週間東洋経済』 6 月 24 日号
- Kohara, M., F. Ohtake, and M. Saito (2002) "A Test of the Full Insurance Hypothesis: The Case of Japan," *Journal of the Japanese and International Economies*, vol.16 (2002) pp.335-352.
- 小宮隆太郎 (1963) 「個人貯蓄の供給」 小宮隆太郎編 『戦後日本の経済成長』 岩波書店
- 松田芳郎・伴金美・美添泰人編 (2000) 『講座ミクロ統計分析 2: ミクロ統計の集計解析と技法』, 日本評論社刊
- 溝口敏行 (1992) 『我が国統計調査の現代的課題』 岩波書店
- 村岸慶應 (1993) 「SNA と家計調査の貯蓄率の比較」 『季刊国民経済計算』 第 99 号 pp.18-79
- 中村洋一 (1999) 『SNA 統計入門』 日本経済新聞社
- 大竹文雄 (2005) 『日本の不平等』 日本経済新聞社
- Ohtake, Fumio and Saito, Makoto, (1998) "Population Aging and Consumption Inequality in Japan," *Review of Income and Wealth* vol. 44(3), pp. 361-81.
- 重川純子 (2007) 「世界の政府機関による家計調査」 御船美智子/家計経済研究所編 『家計研究へのアプローチ』 ミネルヴァ書房 第 3 章
- Shimizutani S. (2006) "Consumer Response to the 1998 Tax Cut: Is a Temporary Tax Cut Effective?" *Journal of the Japanese and International Economies*, vol.20, pp.269-287.
- 清水谷論・堀雅博 (2005) 「ボーナス制度と家計貯蓄率: サーベイデータによる再検証」 『経済研究』 第 56 巻, pp. 234-247.
- 総合研究開発機構 (2008) 『統計改革への提言: 「専門知と経験知の共有化」 を目指して』 NIRA 研究報告書
- 多田吉三 (1989) 『日本家計研究史—わが国における家計調査の成立過程に関する研究』 晃洋書房
- 高山憲之・舟岡史雄・大竹文雄・関口昌彦・澁谷時幸 (1989) 「日本の家計資産と貯蓄率」 『経済分析』 第 116 号 経済企画庁.
- 橋木俊詔・森剛志 (2005) 『日本のお金持ち研究』 日本経済新聞社
- 橋木俊詔・浦川邦夫 (2006) 『日本の貧困研究』 東京大学出版会
- 植田和男・大野正智 (1993) 「家計貯蓄率動向の謎: 世帯調査と国民経済計算との乖離について」 『金融研究 (日本銀行金融研究所)』 第 12 巻第 2 号 127-147 頁
- 梅田雅信・宇都宮浄人 (2003) 『経済統計の活用と論点』 東洋経済新報社

- 宇南山卓 (2006) 「予備的貯蓄とケインズ型消費関数」『社会科学研究 (東京大学社会科学研究所)』第 57 巻第 5・6 合併号
- 宇南山卓 (2008) 「『家計調査』を用いた長期データの作成と応用：パネルデータによる家計消費の分析」総務省統計研修所リサーチペーパー第 10 号
- 宇南山卓 (2009) 「SNA と家計調査における貯蓄率の乖離：日本の貯蓄率の低下の要因」第 10 回マクロコンファレンス (於：一橋大学) 発表論文
- 山田茂 (1990) 「『家計調査』結果の評価に関する一考察」『(国士舘大学) 政経論叢』71
- 谷沢弘毅 (1999) 『経済統計論争の潮流』多賀出版.
- 美添泰人・荒木万寿夫 (2006) 「1980 年代以降における家計の消費と資産に関する実証研究」『総務省統計研修所リサーチペーパー』第 5 号.

家計調査とサンプリング

宇南山 卓 (神戸大学)

家計調査とは？

- ◎ 「国民生活における家計収支の実態を把握し、国の経済政策・社会政策の立案のための基礎資料を提供することを目的」とした統計
 - 政府の公式統計としての利用
 - 消費者物価指数作成のためのウエイト算出
 - 国民経済計算におけるQE推計
 - 景気動向指数
 - 家計の社会経済的分析資料としての活用
 - 景気判断の指標
 - 収入・支出の動向・分布の把握
 - 家計行動の分析
 - 政策評価

家計調査の調査方法

- ◎ 調査世帯数は約9,000世帯
 - 層化三段抽出法による選定
 - 毎月一部が交代するローテーションパネル方式
 - 原則として調査拒否は認められていない
 - 現実には、一定の拒否世帯が存在している
- ◎ 調査項目
 - 世帯員の年齢・性別・就業状態・就学状態
 - 住居の所有形態
 - 毎月の収入(勤労者世帯・無職世帯のみ)・支出
- ◎ 調査方法
 - 自由記入の家計簿記入方式
 - アフターコード方式

調査の母集団世帯の拡大と現状

	二人以上の世帯		単身世帯
	農林漁家を除く	農林漁家を含む	(農林漁家を含む)
勤労者世帯	◎	2000年から	2002年から
無職世帯	1986年から		
その他の世帯	収入は調査せず	収入は調査せず	収入は調査せず

- ◎ 現在調査対象とされていない世帯
 1. 料理飲食店、旅館又は下宿屋を営む併用住宅の世帯
 2. 賄い付きの同居人がいる世帯
 3. 住み込みの営業上の使用人が4人以上いる世帯；
 4. 世帯主が長期間（3か月以上）不在の世帯；
 5. 外国人世帯
 6. その他（単身の学生世帯など）

本日のメニュー

- ◎ 現行の家計調査の問題点の整理
 - 標本誤差
 - 問題の所在とその対応
 - 非標本誤差
 - サンプルセレクションバイアス
 - 測定誤差
 - 調査方法との関係
 - 調査項目・公表項目への要望
- ◎ 収入・支出に関する世帯調査の展望
 - パネルデータ
 - 調査方法の改善

景気指標としての家計調査:標本誤差

- ◎ 家計調査の最大の批判者＝市場エコノミスト
 - 家計調査を景気判断の指標として利用
 - 家計調査の速報性は評価
 - 「毎月の変動が大きすぎる」と批判
 - ◎ 景気判断に求められる「精度」
 - 1%以下の変動によって景気判断が左右される
 - 家計調査の標準誤差率2.0%
- ⇒ たとえ「平均的に見て正しい」としても誤差が大きくては景気指標として役に立たない
- サンプルサイズを増やすことで標本誤差を小さくするべきなのか？

家計の標本誤差とその対応

- ◎ 標本誤差が大きいのは標本数の問題か？
 - 購入頻度の小さな品目は誤差が大きい
 - 標本誤差ではない可能性が高い(=別途対応が必要)
 - そもそも家計消費は実態として不規則
- ◎ 標本数を「十分に」増やすことはできるか？
 - 全国消費実態調査で約6万世帯
 - 家計調査との乖離はそれほど大きくない
 - 現在のサンプル数の400倍が必要(会田,2000)
 - 速報性は維持する必要がある
- ⇒ 他の統計で補完するのが現実的
 - 家計消費状況調査の活用が有効
 - 家計消費状況調査の公表方法の改善が現実的

家計調査の非標本誤差 (1)

- ◎ サンプルセレクションバイアス
 - 日本の「家計全体」を代表したサンプルか？
 - 公務員世帯が多い ⇒ ×
 - 世帯主に限定すれば、1~2%程度多いだけ
 - 公務員世帯の行動は他の世帯とほぼ同じ
 - 専業主婦世帯が多い ⇒ ○
 - 1世帯当たりの有業人員が少ない
 - 次のスライドを参照
 - 低所得世帯・超高所得世帯が少ない ⇒ △
 - 直接比較が可能な統計が限られている
 - 国民生活基礎調査よりは低所得者の割合が低い

家計調査の有業人員

- ◎ 家計調査の有業人員は家計消費状況調査と比較して約0.2人少ない
 - 世帯人員は「比推定」によって調整済み
 - 世帯主以外の有業者を全て「妻」とすれば専業主婦世帯の比率が約20%高いことを意味する
 - 調査負担と関連があると考えられる

パネルC. 二人以上の世帯・勤労者世帯			
	家計調査	全国消費実態調査	家計消費状況調査
集計世帯数	4,427	31,025	10,180
世帯人員 (人)	3.48	3.52	3.49
有業人員 (人)	1.63	1.70	1.82
世帯主の年齢 (歳)	46.4	46.5	48.7

家計調査の非標本誤差 (2)

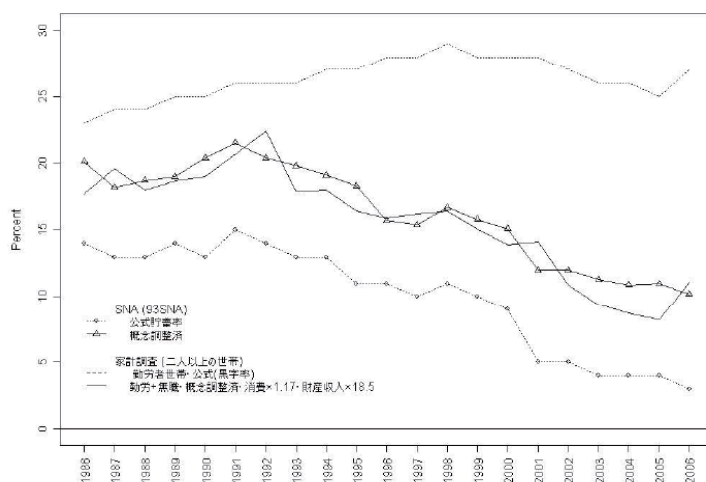
- ◎ 測定誤差
 - 各世帯の支出が正確に把握されているか？
 - 高額消費（耐久消費財・冠婚葬祭費等）が過少⇒ ○
 - 家計消費状況調査と比較して約70%
 - 次のスライドを参照
 - 財産所得が過少 ⇒ ○
 - 1990年の時点で年間3.6万円程度
 - マクロ統計で見れば97万円程度あるはず
 - 家計の金融資産と比較しても過少と考えられる
 - 貯蓄率が低下していない ⇒ ○
 - 岩本・尾崎・前川(1995, 1996)
 - 対象世帯の違いを中心に調整・完全には解明できず
 - 宇南山 (2008)
 - 対象世帯の違い・高額消費・財産所得等の誤差の複合物

高額消費の把握

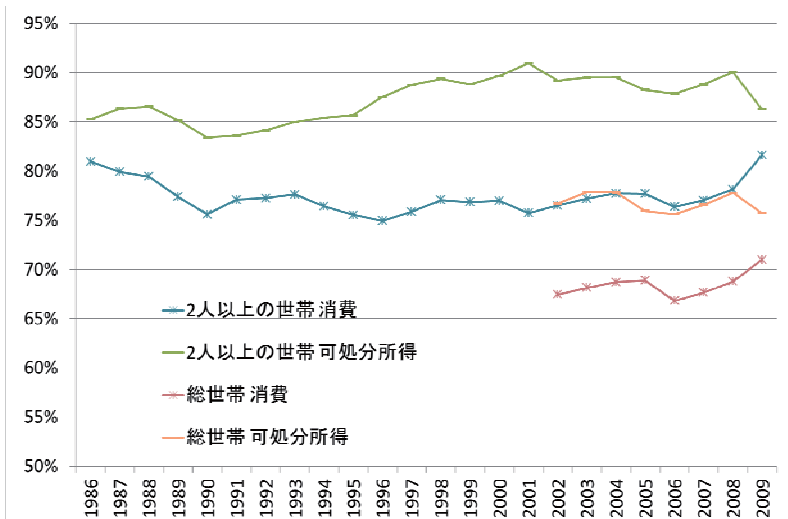
- ◎ 購入頻度が低いため、購入した世帯が調査対象となる可能性が低い
 - 過少となる原因ではないと考えられる
 - サンプル数の大きい全国消費実態調査でも同じ傾向
- ◎ 家計簿に記載すべきかに誤解がある可能性
- ◎ 2002年から家計消費状況調査が開始され、定量的に評価することが可能となった

	2002	2003	2004	2005	2006	2007	2008
	消費支出						
(1) 家計消費状況調査	360,207	361,483	356,243	355,088	358,620	358,450	352,878
(2) 家計調査	305,953	301,841	302,975	300,531	294,943	297,782	296,932
(2)/(1)	84.9%	83.5%	85.0%	84.6%	82.2%	83.1%	84.1%
(1) - (2)	54,254	59,642	53,268	54,557	63,677	60,668	55,946
	調査対象品目の合計						
(3) 家計消費状況調査	148,568	146,296	141,617	142,606	146,252	144,112	140,136
(4) 家計調査	100,819	99,064	101,323	98,928	97,245	98,616	96,456
(4)/(3)	67.9%	67.7%	71.5%	69.4%	66.5%	68.4%	68.8%
(3) - (4)	47,749	47,232	40,294	43,678	49,007	45,496	43,680

SNAと家計調査の貯蓄率



家計調査の「カバー率」



調査方法と非標本誤差

- ◎ 家計簿自由記入・アフターコード方式
 - 消費者物価指数のウェイト算出に必須
 - 詳細な家計簿の記入は調査客体の負担が大きく、負担に耐えられる世帯だけが調査される
 - サンプルセレクションバイアスの原因
 - どのような項目が調査対象であるかの誤解の余地
 - 記入漏れによる測定誤差の原因
- ◎ 家計消費状況調査はプリコード方式
 - 記入すべき項目を事前に指定して金額のみを記入
 - 調査負担は少なく、誤解の余地も小さい
 - 耐久消費財の調査には適している
 - 家計消費状況調査による補完が有効!
 - 本来区分すべきものが含まれる可能性

調査方法改善の可能性

- ◎ アフターコード方式を維持しながら調査負担を軽減する代替的調査方法
 - 記憶による調査
 - 過去に導入が検討されたが、誤差が大きいため採用されてなかった(溝口, 1992)
 - 調査機器の導入による負担軽減
 - 入力支援機器の導入
 - POSシステムによる購入品目の把握
 - プリコード方式との併用
 - 米国のCEXでは記憶によるプリコード方式の調査と家計簿記入方式を併用している
 - プリコード方式には入力支援機器が有効

調査項目・公表項目への要望

- ◎ 追加的な調査項目
 - 学歴・無職の理由・公的支援の有無など
- ◎ 公表項目への希望
 - 年齢別の統計の充実
 - 1歳刻みの年齢別データ
 - コーホートデータの活用が可能になる
 - 年齢と所得・消費水準のクロス集計
 - 消費水準別集計
 - ライフサイクル仮説に基づく「生涯所得」の代理変数
 - 大竹(2005)でも指摘されている
 - 準調査世帯票に関する情報
 - 調査拒否に関する分析が可能
 - 総務省統計局統計図書館で閲覧は可能

収入・支出に関する世帯調査

- ◎ 繰り返しクロスセクション統計としては、家計調査は極めて質の高い統計
 - 米国のCEX・英国のFESと同等以上の貴重な統計
- ◎ パネルデータへの要請
 - 米国のPSID・英国のBHPSなどの長期パネルの試み
 - 経済学における動学的な側面への関心の高まり
- ⇒ 家計調査のパネルデータ化は可能か？
 - 短期パネルとしては原理的には可能
 - 世帯のIDが整備されていないため、容易ではない
 - 長期パネルとするなら調査負担の軽減が必須
 - 他の統計との照合可能性を高める

まとめ

- ◎ 標本誤差を小さくすることは困難
 - 消費は実態として不規則である可能性がある
 - そのまま景気指標として利用するのは困難
 - 家計消費状況調査の活用が有効
 - 家計消費状況調査の公表形態の改善が望まれる
- ◎ 無視できない非標本誤差が存在している
 - 家計簿記入・アフターコード方式が原因
 - 調査方法の改善を検討する余地はある
 - プリコード方式との併用・IT機器の導入
- ◎ パネルデータへのニーズは高まっている
 - 少なくとも短期パネルとして利用できるIDが必要
 - 長期パネルは導入を検討する価値がある

公的統計の開示リスク評価—労働力調査の論点

星野 伸明*

概要

既に提供されている匿名データについて、個体識別の有無は観測可能な事実である。これはデータの匿名化の程度を定める上で、貴重なエビデンスとしなければならない。一般に多様なデータ提供の経験を統一的に把握するには、同一の開示リスク測度で相対評価するのが適当である。そしてこの評価では、ピットマンモデルにより母集団一意数を推定し、開示リスク測度としてファイルに含まれる母集団一意数を用いるのが現実的である。またこのように考えると、労働力調査のローテーション構造を開示リスク評価の改善に使える可能性が見える。

キーワード: 統計的開示制限, 母集団一意, ピットマンモデル.

1 はじめに

個票データは調査客体が知られたいくない情報を「開示 (disclosure)」する可能性がある。このような危険性は、利用者の制限とデータ情報量の削減を組み合わせ管理すれば良い。ここで情報量を削減するようなデータ変換を「匿名化」と呼ぶ。公的統計について言えば、利用者の制限は法的な制度設計の問題である。制度は頻繁には変わらないので、例えば匿名データ作成のような問題では所与の条件として良い。故に本稿では匿名化の評価のみ考察する。

一般に匿名化の評価基準は「開示リスク」及び「有用性」と考えられている。データの情報量を削減すれば、開示の危険性は減るがデータ分析における有用性も減るはずである。従って相反する基準の間でバランスをとらなければならない。方針としては、開示リスクを許容範囲に収めた上で出来るだけ有用なデータを提供すべきである。従って一義的に問題になるのは、開示リスク評価である。

しかし公的統計の匿名化において、個票開示リスク評価の理論は十分活用されていないように思われる。これは日本の実務家向けに処方箋が書かれていない事も一因であろう。故に本稿では匿名化実務での活用を念頭に置き、開示リスク評価の理論を記す。開示リスク概念は役に立たないと捨てられるほど単純でもなく、理屈抜きに使えるほど単純でもない。

匿名データの提供は匿名化が比較的易しい統計調査から難しい統計調査へ移行しつつあり、開示リスク評価について新しい論点が生まれている。例えば5年毎の統計調査の開示リスク評価で

*金沢大学経済学類, 〒 920-0927, 石川県金沢市角間町, E-mail: hoshino@kenroku.kanazawa-u.ac.jp

は時系列構造は無視されるが、このたび匿名データが提供される労働力調査は月次調査である。時系列構造を利用して開示リスク評価を改善できるかもしれない。

本稿では旧統計法の目的外使用（総務省告示第四百八十三号）で得た労働力調査の（月次）基礎票を匿名化¹し、各月独立で評価した開示リスクを報告しておく。結果として開示リスクは自己相関を示す。労働力調査で採用されているローテーションサンプリングでは近接時点の標本が同一個体を含むので、開示リスクの自己相関は構造的と考えられる。

ローテーションサンプリングが広範に使われていることもあり、開示リスクの自己相関はかなり普遍的な現象のはずである。しかし時系列構造に着目した開示リスク評価は、世界的にもほとんど研究されていない。本稿は超母集団モデルの母数に関する時系列モデリングを示唆し、時系列分析専門家の助力を求めるものである。

2 公的統計の開示リスク評価

センシティブ変数を含む個票データの開示リスクを評価するとして、本当に知りたいのはあるレコードが特定の個体と識別される可能性²であろう。しかしこの可能性は、データの情報量以外の様々な要因に依存する。例えば攻撃者が実際に識別を試みるインセンティブの程度、マッチングによる識別を試みる際に使える外部情報の多寡、などが重要である。つまり本当に知りたい事を測るのは、手元のファイルだけ見えても不可能である。

しかしデータの情報量以外の様々な要因は、匿名化の程度を色々変えても不変³とみなして差し支えない。故に匿名化の程度を定める際、データの情報量だけに依存する開示リスク測度を用いる事は合理的である。

では開示リスク測度がデータの情報量だけに依存するとして、いかに構成すれば良いだろうか。これはデータが嘘（「攪乱⁴」）を含むか否かで分けて考えた方が良い。以下では攪乱を含まない場合の開示リスク測度のみ考察するが、これは日本において攪乱されたデータの受容が進んでいないのと、攪乱を含む場合の研究が未熟なためである。

攪乱無しの匿名化技法は、一部のレコードを非公開とする「サブサンプリング」と「再符号化⁵」に事実上限られる。再符号化とは点の観測値を区間表示する事である。区間がもとの観測値を含んでいれば、攪乱とはみなされない。この再符号化は、全レコードに同じ区間分けを適用する「大域的」な場合と、レコード毎に区間分けを変える「局所的」な場合がある。おそらく公的統計では大域的再符号化しか使われないので、以下では局所的再符号化を無視する。局所的再符号化の開示リスク評価については **Hoshino (2009)** が部分的に議論している。

¹本研究の匿名化は匿名データの提供より前に行われており、匿名データの匿名化手法とは異なる。

²識別による開示の他に推測による開示があり得るが、これは匿名データの作成において直接の統制対象ではない。

³センシティブ変数に匿名化を施すと攻撃のインセンティブを低下させるので「不変」ではない。しかし匿名データの作成など一般的な慣行では個体識別リスクの低減が問題とされ、この場合キー変数だけ匿名化すれば十分である。

⁴ここではレコードの情報と対応する個体が正の確率で母集団に存在しない場合を「攪乱された」と言う。詳しくは星野 (2010) を参照のこと。

⁵トップコーディングや削除 (suppression) も再符号化の特殊ケースである。サブサンプリングも局所的な全変数の削除と見なす事が出来るが、このように考えても益は無い。匿名データでは三つ子のレコードを削除していて、局所的匿名化と見なす事が出来るが、やはりサブサンプリングの一種とした方が扱いやすい。

匿名化手法として大域的再符号化のみ使われる場合、個票データセットは情報損失無く分割表に変換可能である。各変数の区間分けが分割表の（周辺）カテゴリ分けに対応しており、全変数の多元分割表から個票データセットを復元出来る。これは単純な事実だが、表章されているか否かで調査票情報とそれ以外を区別するのは無理がある事を示す。また個票データセットの情報から開示リスク測度を構成する際に、分割表を基礎として良い事になる。

分割表とはセルの番地と度数の組の集合である。番地とは、変数の組み合わせで定まる属性空間でのセルの位置を示す。このような情報から開示リスク測度を構成する場合、番地を無視するか否かが第一の問題である。

例えば最も有名な開示リスク測度である「母集団一意数」は、母集団での度数1のセル数であり番地を無視している。もし番地を用いるなら、同じ一意のセルでもその位置により危険性が違うことになる。レコードレベルリスク⁶概念では、このように考える。比較的計算しやすい例としては、より低次元の周辺分割表で一意になるセルの方が危険とみなす“**Minimum Unsafe Combination**”が挙げられる。一般に属性空間の疎な部分に位置する一意のセルは、密な部分の一意のセルよりも識別しやすい。故に番地を開示リスク測度に用いるのは意味があるかもしれない。

しかし番地を利用するには、属性空間で近さを定めなければならない。公的統計の属性変数は名義尺度から比尺度まで入り交じり、適切な近さの設定には個別性の強いモデリングが要求される。実は識別に使えるキー変数の空間のみ考えれば良く、キー変数は様々な統計調査で共通する部分が多い。故に一度近さを定めれば使い回せる可能性は有る。ただ複雑化は避けられない。そして以下で議論するように、複雑化の損失は利益によって補償しきれないだろう。

本節冒頭で述べたように、本当に知りたいのは個体識別の可能性である。そしてこの可能性を評価するには、攻撃者のインセンティブや外部情報の多寡を評価しなければならない。攻撃により金銭的利益を直接得られる可能性から有名になれる可能性まで、リスク要因を顕在化させる事は重要である。また民間の構築する個体情報データベース群について、実情を把握しておく⁷事も重要である。しかしこれらの要因が個体識別につながる可能性は、計量が難しい。最大の困難は識別が観測されないことである。個票の公開で先行する海外でも識別が起きないように匿名化しているので、ほとんど観測されない事象⁸の確率の推定を強いられる。そして識別の例が少ないと、攻撃者のインセンティブや外部情報の存在など、要因毎の効果を十分切り分けられるだけの情報があるとは思えない。

ただ、公的統計の匿名化を今後定める上で、要因毎の効果は必ずしも計量しなくてよい。既に提供されている（匿名）データが識別されていないという事実は観測出来て、これはデータセットの情報量とそれ以外の要因が合わさった上での結果である。そして新しいデータセットについて匿名化の程度を定める際に、データセットの情報量以外の要因は既存の提供データとかなり共通する。したがってデータの情報量のみ依存した開示リスク測度で既に提供されているデータと新しいデータセットを比較すれば、データの情報量以外の要因も含めて個体が識別される可能

⁶局所的匿名化には必要な概念である。なおレコードレベルリスクを集計すると通常の（ファイルレベル）リスクとなる。

⁷外部情報の有無は「匿名データの作成・提供に係るガイドライン」でも特記事項である。

⁸開示制限を失敗して個体識別が可能な例は **Sweeney (2002)** が報告している。

性の大小は判断出来る。言い換えれば、データ情報量のみ依存する開示リスク測度は本当に知りたい事を相対的に評価するものである。

匿名データの提供は、日本人の個体識別の可能性についてデータが無い段階から始められた。この場合の匿名性について、安全方向にマージンを大きく取るのはやむをえない。しかし提供実績が重なれば、マージンを減らしデータの有用性が上がるような道筋をつけなければならない。これを合理的に行うには、データ情報量のみ依存する開示リスク測度を比較し、(識別が起きていなければ) 取るリスクを増やしていけば良い。提供するファイルについて全て同じ開示リスク測度で計り、値を蓄積してゆけば、過去の経験を容易に利用出来る。証拠に基づいて二次利用の環境を改善することが重要である。

上記の方向性で開示リスク測度を用いるとして、具体的に構成してみよう。まず個体情報の詳細さに関して、開示リスク測度は単調性を持つことが望ましい。つまり各変数の区間分けが細かくなれば、より危険な値を示すべきである。一般に区間分けを細かくするとセルは分割され、もとのセルより大きな度数を持つセルが生じる事はない。特に度数 **1** のセルの数は、分割について単調非減少である。故に母集団一意数は分割表の解像度について単調性を持つ。この美点は開示リスクの相対比較をする上で際立つ。母集団一意は論理的な識別可能性を持つことから重要とみなされているようだが、別の良い性質を持つことは強調されて良い。

先に述べたように、母集団一意数は分割表の情報のうち度数しか利用していない。しかし番地を用いるように修正すると、単調性が失われてしまうだろう。また一般に開示リスク測度を微修正するような精緻化は必要無い。それは計量されたリスクの微差が本当に知りたいことについての程度の差を生むか、確定出来るだけの観測情報が期待出来ないからである。

とはいえ母集団一意数は、そのままではファイルの開示リスク測度として使えない。あくまでも母集団の性質を表しているのであって、標本(ファイル)の性質は直接表さないからである。ファイルレベルリスク測度への変換は簡単で、「ファイルに含まれる母集団一意の数」を開示リスク測度とすれば良い。これはファイルのレコード数が n で母集団サイズが \tilde{n} ならば、母集団一意数に n/\tilde{n} をかければ推定出来る。この測度はサブサンプリングの効果も素直に反映する。

関連して注意しておく、世帯単位のファイルでは、本来世帯毎の固まりを「レコード」としてリスク評価すべきである。例えば、世帯員のレコードを年齢順に連結したまとまりを一レコードとして扱えば良い。この場合、世帯人数が異なればレコード長も異なる。しかし相対評価が前提ならば、便宜的に一個体を一レコードとして評価しても良いだろう。このようにすれば、(個人単位でサブサンプリングするなどして) 世帯構造を崩して公開されるファイルとの比較も可能となる。

残る問題は母集団一意数の推定である。これが難しい理由はかつて説明した(星野, 2003)。一点つけ加えておくと、母集団一意数の推定改善にセルの番地情報を使うアプローチはあり得る。しかし大規模かつ疎な分割表では絶対的に情報が不足しているので、うまくいかないであろう。またそのようなアプローチは高度なモデリングが要求され、開示リスク評価の試行錯誤にも向かない。故に実務への採用は難しいかもしれない。以下では最も現実的と考えられる推定方法とその

妥当性を説明する。

幅広い母集団について推定精度をルーチンワークとして確保するには、ピットマンモデルを用いるのが良い。このアプローチにおいてデータは、無限母集団すなわち **Pitman** 分布からの標本とみなされる。そして母集団⁹も同一無限母集団からの標本とみなすので、データから **Pitman** 分布の母数を最尤推定し、推定値の下で母集団一意数の挙動を求める。より具体的には、付録の手順書を参照されたい。また **Pitman** 分布は 2.1 節で紹介される。

大事な事は同じ手法で計り続ける事である。開示リスクを評価するファイルのレコード数は、母集団個体数のせいぜい一割程度であろう。この場合にある程度安定的な母集団一意数の推定量は、全てバイアス¹⁰を持つ。手順書の推定量も例外でなく、おそらく過大に一意数を推定する。しかし開示リスクを相対比較するのが目的なので、バイアスは方向がそろっている限り、あまり問題にならない。このように考えると、常に同一手法で母集団一意数を推定するべきである。

なお特定のモデルと決めつけるよりも、モデル集合からデータに良く当てはまるモデルを選択し、そのモデルで一意数を推定する方が正確になろう。しかしモデル集合の空間をうまく張らないと、匿名化の試行錯誤においてリスク評価値がぶれる。また経験的に多くの場合、ピットマンモデルが選択¹¹される。故に手間や精度及び様々な結果の整合性を勘案すれば、母集団一意数は常にピットマンモデルによって推定するのが最善と思われる。

2.1 ピットマンモデルについて

本節では数理に踏み込み、ピットマンモデルによる推定の方法論を理解する。単に母集団一意数を推定するだけなら、付録の手順書に従えばよい。

自然数 $n \in \mathbb{N} := \{1, 2, 3, \dots\}$ を自然数の和で表す事を分割と呼ぶ。この和の中で自然数 i が足される回数を s_i で表せば、 $\mathbf{s}_n := (s_1, s_2, \dots, s_n)$ は (順序無し) の分割を表す。非負整数の集合を \mathbb{N}_0 で表すと、 n の全ての分割の集合は $\mathcal{S}_n := \{\mathbf{s}_n : s_i \in \mathbb{N}_0, i \in \{1, 2, \dots, n\}, \sum_{i=1}^n i s_i = n\}$ で表される。この集合上の分布が自然数の確率分割である。以下では $u := \sum_{i=1}^n s_i$ とする。

Pitman 分布は自然数の確率分割であり、母数 $0 \leq \alpha < 1, \theta > -\alpha$ について確率関数は以下のように書ける。

$$p(\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n) = n! \frac{\theta^{[u:\alpha]}}{\theta^{[n]}} \prod_{j=1}^n \left(\frac{(1-\alpha)^{[j-1]}}{j!} \right)^{s_j} \frac{1}{s_j!}, \quad \mathbf{s}_n \in \mathcal{S}_n, \quad (1)$$

⁹一部が観測されている現実の母集団について推定するのではなく、同サイズの母集団を新たに発生させる場合の挙動が推定される。

¹⁰有限母集団から非復元単純無作為抽出する場合、一意数の不偏推定量は一意に存在する。しかしこの不偏推定量は標準誤差が大きく、標本抽出率がかかり高く (8割程度と思う) ないと実用に耐えない。そして一意な不偏推定量なので、推定を安定させるためのいかなる工夫もバイアスを生む。

¹¹裾の長いモデルとして代表的な負の二項分布は、統計的開示制限の分野ではポアソン=ガンマモデルとして知られている。このモデルは基本的に広義のピットマンモデルの特殊ケース ($\alpha \leq 0$ に対応) である。故にピットマンモデルのデータへのあてはまりは、基本的にポアソン=ガンマモデルを下回らない。そしてポアソン=ガンマモデルによる母集団一意数の推定値は、必ず **Pitman** モデルの推定値より (かなり) 小さくなると考えて良い。

ただし $\theta^{[u;\alpha]} = \theta(\theta + \alpha) \cdots (\theta + (u-1)\alpha)$, $\theta^{[n]} = \theta(\theta + 1) \cdots (\theta + n - 1)$ である。

個票データとの対応を述べておこう。匿名化の程度を定めることで、キー変数に関する分割表が出来る。分割表の情報のうち度数を、第 j セルについて f_j と書く。ただしセル総数が J として $j \in \{1, 2, \dots, J\}$ である。ここで $i = 1, 2, \dots, n$ について度数 i のセルの数を s_i と表す。つまり指示関数 $1(\cdot)$ を使えば、 $s_i = \sum_{j=1}^J 1(f_j = i)$ である。例えば s_1 は標本で一意なレコード数となる。このように作られる s_n を「寸法指標」と呼び、 $n \ll J$ なら Pitman 分布の標本とみなせる。母集団サイズ \tilde{n} も J よりかなり小さいなら、確率変数 $\mathbf{S}_{\tilde{n}} := (S_1, S_2, \dots, S_{\tilde{n}})$ が Pitman 分布に従う場合、 S_1 で母集団一意数の挙動が表せる。 $\tilde{n} \ll J$ でない場合の対応は付録で言及している。

経験ベイズ的に母集団一意数推定の論理を説明すると、以下の通りになる。 $\mathbf{S}_{\tilde{n}}$ の事前分布が Pitman 分布であり、その母数 (α, θ) は超母数である。超母数はデータ s_n により (最尤) 推定される。推定したい母数は S_1 であり、 $(S_2, S_3, \dots, S_{\tilde{n}})$ は局外母数である。このように考えれば母数 S_1 の周辺分布が必要となる。 $U_n := \sum_{i=1}^n S_i$ として、証明抜きで一般の結果を述べる、

Theorem 1 *Let (\mathbf{S}_n, U_n) be Pitman distributed as (1). Then*

$$\Pr(S_i = s, U_n = u) = \frac{n! \theta^{[u;\alpha]} \sum_{j=0}^{u-s} \frac{(-1)^{u-s-j} \alpha^{-j} |C(n-i(u-j), j, \alpha)|}{s! \theta^{[n]} (n-i(u-j))! (u-s-j)!}, \quad (2)$$

$$s \in \{0, 1, \dots, \lfloor n/i \rfloor\},$$

$$u \in \{s, s+1, \dots, s + \lfloor (n-s)/2 \rfloor\} \quad \text{if } i = 1,$$

$$\{s, s+1, \dots, s + (n-is)\} \quad \text{if } i \in \{2, 3, \dots, n\}.$$

定理 1 で $C(\cdot, \cdot, \cdot)$ は C-ナンバーと呼ばれ、一般化されたスターリング数である。C-ナンバーについては Charalambides and Sing (1988) を参照のこと。 S_i の周辺確率は (2) を全ての u について足せば求められる。

Pitman 分布に従う \mathbf{S}_n の任意の周辺階乗モメントは、Yamato and Sibuya (2000) が与えている。すなわち非負整数の $r_i, i = 1, 2, \dots, n$, について

$$\mathbb{E}\left(\prod_{i=1}^n S_i^{(r_i)}\right) = \frac{\theta^{[r;\alpha]} (\theta + r\alpha)^{[n-R]} n^{(R)}}{\theta^{[n]}} \prod_{i=1}^n \left(\frac{(1-\alpha)^{[i-1]}}{i!}\right)^{r_i},$$

ただし $r = r_1 + \dots + r_n, R = \sum_{i=1}^n i r_i \leq n$ である。特に

$$\mathbb{E}(S_i) = \frac{(1-\alpha)^{[i-1]} n^{(i)}}{i!} \theta \left(\frac{(\theta + \alpha)^{[n-i]}}{\theta^{[n]}}\right), \quad (3)$$

$$\begin{aligned}
V(S_i) &= \left(\frac{(1-\alpha)^{[i-1]}}{i!}\right)^2 \theta^{[2:\alpha]} \left(\frac{(\theta+2\alpha)^{[N-2i]} n^{(2i)}}{\theta^{[n]}}\right) \\
&\quad + \frac{(1-\alpha)^{[i-1]} n^{(i)}}{i!} \theta \left(\frac{(\theta+\alpha)^{[n-i]}}{\theta^{[n]}}\right) \\
&\quad - \left(\frac{(1-\alpha)^{[i-1]} n^{(i)}}{i!} \theta \left(\frac{(\theta+\alpha)^{[n-i]}}{\theta^{[n]}}\right)\right)^2
\end{aligned}$$

となる。

以上の結果について $n = \tilde{n}, i = 1$ とすれば、母集団一意数の挙動となる。超母数の最尤推定量を $(\hat{\alpha}, \hat{\theta})$ と書けば、(3) 式に $\alpha = \hat{\alpha}, \theta = \hat{\theta}$ を代入して母集団一意数の点推定量が得られる。すなわち

$$\hat{S}_1 = \tilde{n} \frac{(\hat{\theta} + \hat{\alpha})(\hat{\theta} + \hat{\alpha} + 1) \cdots (\hat{\theta} + \hat{\alpha} + \tilde{n} - 2)}{(\hat{\theta} + 1)(\hat{\theta} + 2) \cdots (\hat{\theta} + \tilde{n} - 1)}. \quad (4)$$

なお Hoshino (2001, Proposition 3) によれば、 $\alpha \geq 0$ について $\lim_{n \rightarrow \infty} E(S_1)/E(U_n) = \alpha$ である。 U_n は度数が 0 でないセルの総数なので、母集団で空でないセルのうち一意のセル数の割合は α と解釈出来る。このように Pitman 分布の母数 α は、開示リスクと直結する解釈を持つ。

従って超母数、特に α を開示リスク測度とみなすことは自然である。最尤推定量 $(\hat{\alpha}, \hat{\theta})$ は、開示リスクを直接推定していると考えてもよい。この場合に (α, θ) の推定誤差の重要性は増す。フィッシャー情報量を確認しておこう。まず対数尤度関数を

$$L(\alpha, \theta) = \sum_{i=1}^{u-1} \log(\theta + i\alpha) - \sum_{i=1}^{n-1} \log(\theta + i) + s_1 + \sum_{i=2}^n s_i \sum_{j=1}^{i-1} \log(j - \alpha) + \text{Const}. \quad (5)$$

で表す。二次の微分係数は

$$\frac{\partial^2 L(\alpha, \theta)}{\partial \theta^2} = - \sum_{i=1}^{u-1} \frac{1}{(\theta + i\alpha)^2} + \sum_{i=1}^{n-1} \frac{1}{(\theta + i)^2}, \quad (6)$$

$$\frac{\partial^2 L(\alpha, \theta)}{\partial \alpha^2} = - \sum_{i=1}^{u-1} \frac{i^2}{(\theta + i\alpha)^2} - \sum_{i=2}^n s_i \sum_{j=1}^{i-1} \frac{1}{(j - \alpha)^2} < 0, \quad (7)$$

$$\frac{\partial^2 L(\alpha, \theta)}{\partial \theta \partial \alpha} = - \sum_{i=1}^{u-1} \frac{i}{(i\alpha + \theta)^2} < 0 \quad (8)$$

である。(8) 式より $\hat{\alpha}$ と $\hat{\theta}$ は負の相関を持つ。情報量はこれらの式について u を U_n に、 s_i を S_i に置き換えて期待値をとる。 $E(S_i)$ は (3) 式で与えられているので、あとは

$$P(U_n = u) = \frac{\theta^{[u:\alpha]}}{\theta^{[n]}} (-1)^{n-u} C(n, u, \alpha) \alpha^{-u}, \quad u \in \{1, 2, \dots, n\}. \quad (9)$$

を利用して数値的に評価できる。Sibuya and Yamato (2001, Proposition 5) がフィッシャー情報量行列のオーダーを評価しており、 $n \rightarrow \infty$ の時 $I_{\theta\theta} = O(1), I_{\theta\alpha} = O(\log n), I_{\alpha\alpha} = O(n^\alpha)$ である。特に θ の推定精度は悪い。

最尤推定は母数変換について不変¹²である。 (α, θ) という開示リスク測度の関数として書ける任意の開示リスク測度について、 $(\hat{\alpha}, \hat{\theta})$ を代入すれば最尤推定量となる。例えば (3) 式の右辺は (α, θ) の関数であり、一次元に退化するように変数変換した開示リスク測度とみなせる。つまり母集団一意の点推定量 \hat{S}_1 も、開示リスク測度 $E(S_1)$ を最尤推定している。このように最尤推定は開示リスク測度変換について閉じており、開示リスク測度の選択につきまとう恣意性は緩和される。

3 労働力調査について

ここまで考察したことは、異なるファイルの開示リスクを比較する方法論であった。暗黙の前提として各ファイルは独立である。しかし労働力調査の月次のファイルに同一基準で匿名化を施すとして、開示リスクに無視出来ない自己相関が存在するかもしれない。そして労働力調査は、構造的に自己相関を持つような標本設計「ローテーションサンプリング」が採用されている。

労働力調査の標本設計を確認しておこう。総務省統計局 (2011) によれば、労働力調査は層化二段抽出を用いている。国勢調査の約 100 万調査区を特性により層化し、11 地域別に約 2900 調査区を一次抽出する。二次抽出で住戸を抽出し、約 4 万世帯（基礎調査票）及びその世帯員が調査対象となる。就業状態は 15 歳以上の者について調査している。ローテーションは以下のように行われる。まず一次抽出された調査区は、4ヶ月継続して調査される。この期間の終了後、他の標本調査区に交代するが、翌年同期に再び調査される。二次抽出された住戸は、同調査区で調査される 4ヶ月の前半と後半で入れ替えられる。世帯単位で見れば 2ヶ月連続で調査され、10ヶ月空けた後、再び 2ヶ月連続で調査される。分かりにくいのが、表 1 は各月に調査される世帯がローテーションのどの時期にあたるかを示している。例えば A-1 グループは一月に初めて調査される世帯であり、A-2 グループでは一月に 2 年目の 1 回目調査が行われる。B-1 グループは一月に 1 年目の 2 回目調査が行われ、二月には新しい調査区に移る。

調査区は月毎に全体の四分の一ずつ交代し、世帯は月毎に全体の二分の一ずつ交代する。連続する 2ヶ月で同一人物のキー変数が変化する事は少ない。また近接して住む人間は似るということは、地理学の基本命題だそうだ。開示リスクが自己相関を持つことは十分考えられる。

ここでパネルデータの開示リスク評価について述べておく。労働力調査は 2ヶ月+2ヶ月のパネル構造を持つ。故に匿名化の方法によっては、パネルデータとして開示リスクを評価しなくてはならない。一般に T 期間のパネルデータは、キー変数を T 倍にして開示リスクを評価する。これは分割表で考えると理解しやすい。第一期のキー変数から作られる分割表の各セルは、更に第二期の（同一）キー変数で分割される。ここで出来たセルは三期以降も同様に分割される。つまりクロスセクションで見て J 個のセルは、パネルでは J^T 個になる。パネルの期間が長くなる場合に

¹²例えば竹村 (2001, p.139) を見よ。

Group	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
A-1	1-1	1-2	1-1	1-2]	1-1	1-2	1-1	1-2]	1-1	1-2	1-1	1-2]
A-2	2-1	2-2	2-1	2-2]	2-1	2-2	2-1	2-2]	2-1	2-2	2-1	2-2]
B-1	1-2]	1-1	1-2	1-1	1-2]	1-1	1-2	1-1	1-2]	1-1	1-2	1-1
B-2	2-2]	2-1	2-2	2-1	2-2]	2-1	2-2	2-1	2-2]	2-1	2-2	2-1
C-1	1-1	1-2]	1-1	1-2	1-1	1-2]	1-1	1-2	1-1	1-2]	1-1	1-2
C-2	2-1	2-2]	2-1	2-2	2-1	2-2]	2-1	2-2	2-1	2-2]	2-1	2-2
D-1	1-2	1-1	1-2]	1-1	1-2	1-1	1-2]	1-1	1-2	1-1	1-2]	1-1
D-2	2-2	2-1	2-2]	2-1	2-2	2-1	2-2]	2-1	2-2	2-1	2-2]	2-1

Note: “2-1” implies 2nd year’s 1st month. “[” implies enumeration districts’ change.

表 1: The rotation pattern of the Japanese Labor Force Survey

セル数はかなりのオーダーで増加するが、属性が時系列であまり変化しなければ、一意数はあまり増えない。要するに属性の変化というイベント情報は、個体の識別力を持つということである。

労働力調査の匿名データのようにパネルとして再構成出来ない¹³ なら、時点間の不完全な従属性を自己相関のようにモデル化した方が良くもしいない。

時点間の従属性は、超母集団について同定すべきである。標本（ファイル）の開示リスクの比較が最終的な目標なので、標本開示リスクの従属性を直接モデル化するのが自然かもしれない。しかし標本開示リスクは、標本サイズと母集団サイズの変化の影響をうける。例えば図 1 は、2002 年 1 月から 2007 年 12 月まで労働力調査の標本サイズを示している。このような変化は標本開示リスク評価において無視出来ない。そもそも母集団及び標本のサイズは既知¹⁴と考えられる。これらに依存しない未知な要素、すなわち超母集団の開示リスク構造のみにモデル化の対象を限定する方が良い。また超母集団の開示リスク構造が特定されていれば、標本の開示リスク評価は容易である。

超母集団の開示リスクについて、時点間の従属性の有無を実データで確認しよう。以下では労働力調査の基礎票について、2002 年 1 月から 2007 年 12 月まで月次の開示リスクを検討する。これまでの議論より、各ファイルを個人単位とみなし、各月独立にピットマンモデルの母数の最尤推定値を求める。

各月について 8 通りの匿名化パターンを適用し、匿名化手法の効果も計量する。表 2 は各ケースのキー変数とそのカテゴリー数をまとめた¹⁵ものである。下線部は前のケースから変更されたことを意味しており、その他の変数のカテゴリー数は変化しないので省略してある。つまりケース 1 から単調に匿名化を強くした。ケース 2 では産業コードを削除した。ケース 3 では地域区分を 6(北海道+東北, 関東, 中部, 近畿, 中国+四国, 九州) に再符号化した。ケース 4 では 15 歳未満男

¹³前月の同属性の個体を照合して、統計的なパネルデータは構成できる。この場合の開示リスク評価は攪乱データのそれに準じる。

¹⁴特に母集団サイズは人口学的に推計する方が正確なはずである。

¹⁵2002 年 1 月のファイルに存在するカテゴリーを数えた結果なので、他の時期のファイルでは多少異なるかもしれない。

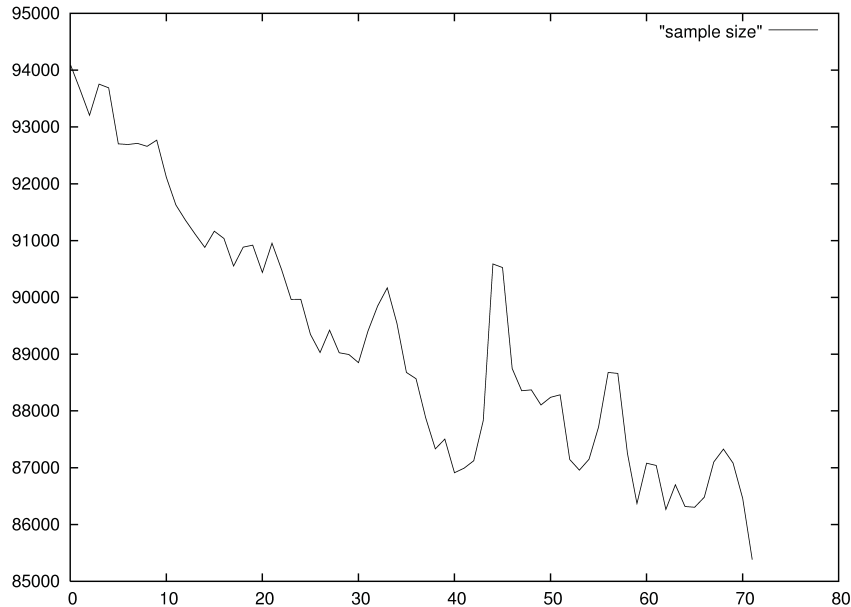


図 1: Sample sizes from 1/02 to 12/07

女数を 3人以上のようにトップコーディングした。ケース 5 では 15 歳以上総数を 5人以上のようにトップコーディングした。ケース 6 では続柄について、祖父母、兄弟姉妹、他の親族、その他、不詳を一つのカテゴリーに再符号化した。ケース 7 では 85 歳以上のようにトップコーディングした。ケース 8 では地域区分（都道府県番号）を削除した。

図 2 はケース 3 から 6 について、72 時点の最尤推定値 $(\hat{\alpha}, \hat{\theta})$ の散布図である。最尤推定量の負の相関が確認できる。また匿名化が加わるにつれ、点が単調に左（上）に移動している。これは開示リスク測度として望ましい性質である。

各ケースを比較しやすいように、72 点の“中心”座標を求める。72 ファイルを全て同一の超母集団からの標本とみなし、その母数 (α^*, θ^*) を最尤推定した結果が表 3 である。ケース 2,3,8 で加えた匿名化が特に効いていることが、母数の変化として表れている。ケース 4 から 7 はケース 3 の匿名化を微修正したものなので、母数も変化が小さい。なお匿名化が強化されると (α^*, θ^*) は全体的に左上に変化しているが、ケース 6 から 7 ではわずかに右下にずれている。

次に t 期の最尤推定値を $(\hat{\alpha}_t, \hat{\theta}_t)$ と見て、時系列変化を確認する。図 3 は 1 から 72 期まで $(t, \hat{\alpha}_t)$ を 8 ケースについてプロットした。同期間に超母集団の開示リスクは増大しているはずで、そのような増加トレンドが見られる。佐井至道教授の指摘のとおり、労働力調査では失業者に聞かない変数がある。故に失業率が低下すれば、個人情報により詳細になり開示リスクは増加する。図 4 は同時期の季節調整されていない失業率を示しており、右下がりのトレンドが見られる。

ケース 1 について $\hat{\alpha}_t$ の自己相関係数（コログラム）を図 5 に示した。時差 1 と 12 のピークが目立つ。時差 12 の自己相関はこれほどはっきりしないケースもあったが、時差 1 の自己相関は全ケースで目立つ。各月独立ではローテーションの情報を使い切っていないということになる。

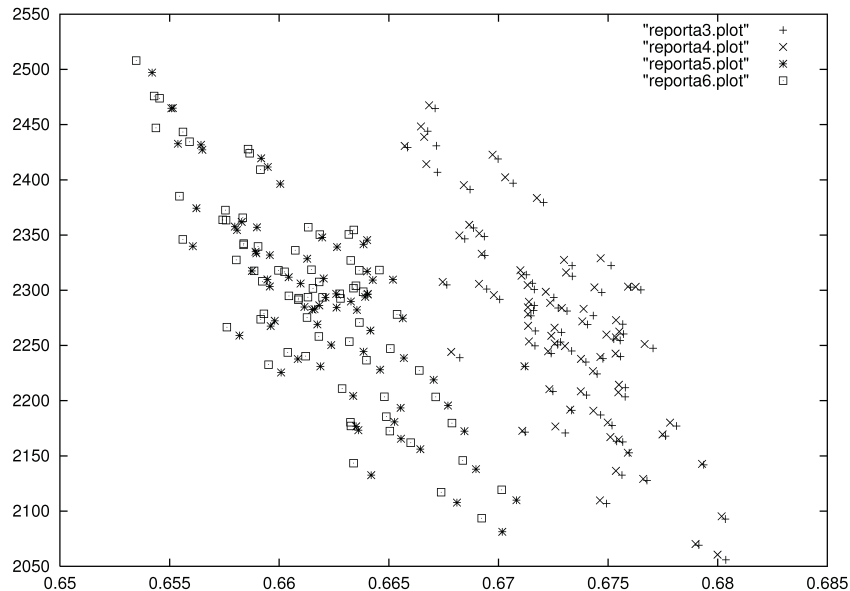


図 2: The plots of $(\hat{\alpha}, \hat{\theta})$ in Case 3-6

ここまで労働力調査の開示リスク評価について、時系列モデリングの必要性を示した。世界的に見ても、既存研究では時間的従属性を利用出来ていない。海外の研究ではクロスセクションのモデリングが複雑で、時系列構造を組み込むのは困難と思われる。一方、超母集団モデルは少数の母数でクロスセクションの開示リスクを表すので、時間依存を利用しやすい。幸い日本の時系列研究者は層が厚いので、協力して重要な貢献が出来るかと期待している。

付録：母集団一意数の推定手順

以下では標本サイズを n 、母集団サイズを \tilde{n} と記す。

1. 評価するキー変数とその精度を決める。
2. 決められたキー変数全てについてクロス集計する。つまり（高次元の）分割表を作り、各セルに所属するレコード数（度数）を数える。
 - 第 j セルの度数を f_j と書く。セル総数が J として $j = 1, 2, \dots, J$ だが、インデクスの順序はいつでも良い。
 - セル総数 J は、全てのキー変数のカテゴリー数の積である。連続変数であっても、開示リスク評価においては（攻撃者の持つ情報精度にあわせて）離散として扱うのが妥当である。
3. 空でないセルの度数の度数（寸法指標）を数える。

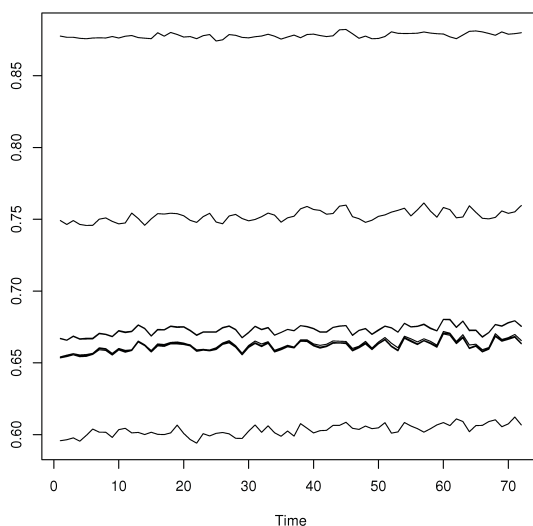


図 3: $\hat{\alpha}_t$ of Case 1-8

- $i = 1, 2, \dots, n$ について度数 i のセルの数を s_i と表す。つまり指示関数 $1(\cdot)$ を使えば、 $s_i = \sum_{j=1}^J 1(f_j = i)$ である。
- 最大のセルの度数が m ならば、 $m < i$ について $s_i = 0$ である。

4. データを生成した構造（確率分布）を推定する。

- 現実の母集団を無限母集団（超母集団）からの標本とみなす。この場合、手元の標本から超母集団の分布を推定すれば、母集団の挙動も推定される。
- 超母集団の分布として広義の Pitman モデルを仮定し、その母数を最尤推定する。
- Pitman モデルは 2 母数 (α, θ) を持ち、 α が負の場合と正の場合で分けて考えた方がよい。どちらの場合も $u = \sum_{i=1}^n s_i$, $n = \sum_{i=1}^n i s_i$ である。

(a) $0 \leq \alpha < 1, \theta > -\alpha$ について Pitman モデルの確率関数は (1) 式で表される。

(b) $\alpha < 0$ の場合は (1) 式で $\theta = -J\alpha$ とおき、さらに $-\alpha = \gamma$ とおく。すると一母数の確率関数を得る：

$$p(s_1, s_2, \dots, s_n) = \frac{n! J! \Gamma(J\gamma)}{\Gamma(J\gamma + n)} \prod_{i=0}^n \left(\frac{\Gamma(\gamma + i)}{\Gamma(\gamma) i!} \right)^{s_i} \frac{1}{s_i!} \quad (10)$$

ここで $\gamma > 0$ であり、 $s_0 = J - u$ である。



図 4: Unadjusted unemployment rate (01/02-12/07)

- モデル (1) を「(狭義の)Pitman モデル」と呼ぶ。モデル (10) を「多項ディリクレモデル」と呼ぶ。本来は AIC 等によりデータ依存でいずれかをモデル選択するのが良いが、ここでは簡易的な選択基準を示す：
 - (a) 母集団サイズ \hat{n} が総セル数 J より大の場合、多項ディリクレモデルを用いる。
 - (b) その他の場合は Pitman モデルを用いるが、尤度の最大化に失敗する（繰り返し計算が収束しない）場合、多項ディリクレモデルを用いる。
- 狭義の Pitman モデルの最尤推定は以下のように行えば良い。
 - (a) 対数尤度関数は (5) 式で与えられている。
 - (b) 最尤推定量は以下の同時方程式の解である。

$$\frac{\partial L(\alpha, \theta)}{\partial \theta} = \sum_{i=1}^{u-1} \frac{1}{\theta + i\alpha} - \sum_{i=1}^{n-1} \frac{1}{\theta + i} = 0,$$

$$\frac{\partial L(\alpha, \theta)}{\partial \alpha} = \sum_{i=1}^{u-1} \frac{i}{\theta + i\alpha} - \sum_{i=2}^n s_i \sum_{j=1}^{i-1} \frac{1}{j - \alpha} = 0.$$

- (c) $L(\alpha, \theta)$ の最大化は汎用最大化ルーチン (R の `optim()` 関数等) に任せても良いだろう。
- (d) 最尤推定値を自前で評価するなら、二次の微分係数 (6),(7),(8) 式を用いたニュートン＝ラフソン法が適当である。

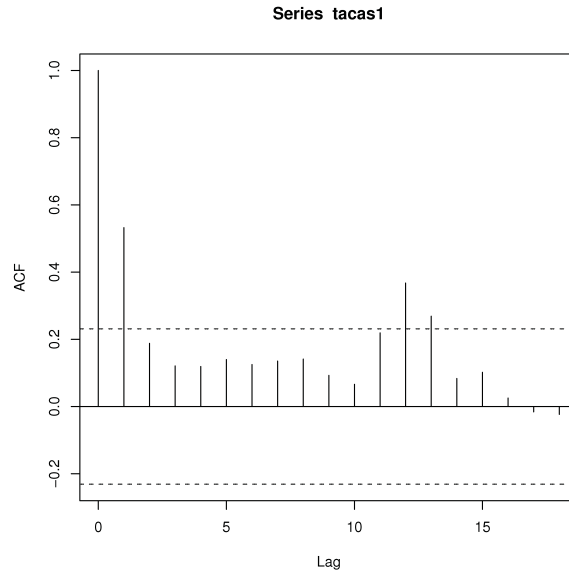


図 5: ACF of $\hat{\alpha}_t$ in Case 1

(e) 近似的モメント推定量 :

$$\hat{\theta} = \frac{nuc - s_1(n-1)(2u+c)}{2s_1u + s_1c - nc}, \quad \hat{\alpha} = \frac{\hat{\theta}(s_1 - n) + (n-1)s_1}{nu},$$

ただし $c = s_1(s_1 - 1)/s_2$ は、ニュートン=ラフソン法の初期値として使えない事もない。

(f) θ の推定は不安定なので、初期値をランダムに変えて繰り返し計算が同じ値に収束するか確認するのが望ましい。

- 多項ディリクレモデルの最尤推定は以下のように行えば良い。

(a) 対数尤度関数は定数を除いて

$$L(\gamma) = -\sum_{i=0}^{n-1} \log(J\gamma + i) + \sum_{i=1}^n s_i \sum_{j=0}^{i-1} \log(\gamma + j).$$

(b) 最尤推定値は尤度方程式

$$\frac{dL(\gamma)}{d\gamma} = -\sum_{i=0}^{n-1} \frac{J}{J\gamma + i} + \sum_{i=1}^n s_i \sum_{j=0}^{i-1} \frac{1}{\gamma + j} = 0$$

の解である。

- (c) $L(\gamma)$ の最大化は汎用最大化ルーチン (R の `optimize()` 関数等) に任せても良いだろう。
- (d) 最尤推定値を自前で評価するなら、二次の微分係数

$$\frac{d^2 L(\gamma)}{d\gamma^2} = \sum_{i=0}^{n-1} \frac{J^2}{(J\gamma + i)^2} - \sum_{i=1}^n s_i \sum_{j=0}^{i-1} \frac{1}{(\gamma + j)^2}$$

を用いたニュートン=ラフソン法が適当である。

- (e) 尤度関数は単峰であり、それほど初期値に依存せず最大化が可能である。ただ最尤推定値が無限大に発散する事はあり得て、それは確率関数が等確率 J 項分布である事を意味する。また最尤推定値が 0 の場合、狭義の **Pitman** モデルの方が適切と思われる。
- 狭義の **Pitman** モデルと多項ディリクレモデルの境界 ($\alpha = 0$) のモデルを **Ewens** モデルという。Ewens モデルの確率関数は以下の通り：

$$p(s_1, s_2, \dots, s_n) = n! \frac{\theta^n}{\theta^{[n]}} \prod_{j=1}^n \left(\frac{1}{j}\right)^{s_j} \frac{1}{s_j!}. \quad (11)$$

- 同じデータについて **Ewens** モデルの最尤推定値を $\hat{\theta}_E$ と書き、**Pitman** モデルの最尤推定値を $\hat{\alpha}, \hat{\theta}_P$ と書く。もし $\hat{\alpha} > 0$ ならば

$$\hat{\theta}_E > \hat{\theta}_P.$$

- 上の式は **Pitman** モデルのチェックに使える。また最尤推定の繰り返し計算の範囲を限定できる。
- **Ewens** モデルの尤度関数は単峰であり、最大化は容易である。

5. 同定されたデータ構造の下で母集団一意数の推定値 \hat{S}_1 を求める。

- (a) 狭義の **Pitman** モデルの場合、母数の最尤推定値を $\hat{\alpha}, \hat{\theta}$ と書けば (4) で推定される。
- (b) 多項ディリクレモデルの場合、母数の最尤推定値を $\hat{\gamma}$ と書けば

$$\hat{S}_1 = \hat{n}(J-1)\hat{\gamma} \frac{((J-1)\hat{\gamma}+1)((J-1)\hat{\gamma}+2)\cdots((J-1)\hat{\gamma}+\hat{n}-2)}{(J\hat{\gamma}+1)(J\hat{\gamma}+2)\cdots(J\hat{\gamma}+\hat{n}-1)}.$$

- これらの推定値はモデルの下での度数 1 のセル数の期待値である。
- 注1) **Ewens** モデルの母集団一意数推定式は、**Pitman** モデルの推定式に $\hat{\alpha} = 0$ を代入して得られる。
- 注2) 等確率 J 項分布の母集団一意数推定値は $\hat{n}(1 - 1/J)^{\hat{n}-1}$ である。

参考文献

- [1] Charalambides, C.A. and Singh, J. (1988) A review of the Stirling numbers, their generalizations and statistical applications. *Communications in Statistics, Theor. Meth.*, **17**, 2533–2595.
- [2] Hoshino, N. (2001) Applying Pitman's sampling formula to microdata disclosure risk assessment, *Journal of Official Statistics*, **17**, 499–520.
- [3] 星野伸明 (2003) 「超母集団モデルによる個票開示リスク評価」, *統計数理*, **51**, 297–319.
- [4] Hoshino, N. (2009) The Quasi-multinomial Distribution as a Tool for Disclosure Risk Assessment, *Journal of Official Statistics*, **25**, 269–291.
- [5] 星野伸明 (2010) 「公的統計マイクロデータ提供制度の課題」, *日本統計学会誌*, **40**, 23–45.
- [6] Sibuya, M. and Yamato, H. (2001) Pitman's model of random partitions. *数理解析研究所講究録*, **1240**, 64–73.
- [7] 総務省統計局 (2011) 『労働力調査標本設計の解説 (平成 20 年 4 月)』
- [8] Sweeney, L. (2002) k -Anonymity: a model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems*, **10**, 557–570.
- [9] 竹村彰通 (1991) 『現代数理統計学』, 創文社, 東京.
- [10] Yamato, H. and Sibuya, M. (2000). Moments of some statistics of Pitman sampling formula. *Bulletin of Informatics and Cybernetics*, **32**, 1–10.

Key Variable	Case 1	Case 2	Case 3	Case 4
Prefecture code	47		<u>6</u>	
Persons 15 years old & over	9			
– under 15 years old (male)	5			<u>4</u>
– under 15 years old (female)	5			<u>4</u>
Sex	2			
Relationship to the head	12			
Age	19			
Marital status	4			
Employment status	9			
Hours worked	3			
Occupation	9			
Organizational form	4			
Industry	79	<u>1</u>		
Number of persons engaged in enterprise	5			

Key Variable	Case 5	Case 6	Case 7	Case 8
Prefecture code	6			<u>1</u>
Persons 15 years old & over	<u>5</u>			
– under 15 years old (male)	4			
– under 15 years old (female)	4			
Sex	2			
Relationship to the head	12	<u>8</u>		
Age	19		<u>15</u>	
Marital status	4			
Employment status	9			
Hours worked	3			
Occupation	9			
Organizational form	4			
Industry	1			
Number of persons engaged in enterprise	5			

表 2: The number of categories

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8
α^*	.878	.752	.673	.673	.662	.661	.663	.603
θ^*	2608.7	7041.7	2260.1	2263.4	2287.0	2295.3	2224.5	6554.5

表 3: The center of the 72 months

パネルデータの作成の実情と課題

北村行伸

一橋大学経済研究所

1. はじめに

日本では 1994 年度より家計経済研究所によって「消費生活に関するパネル調査」が始まり、2000 年代には 21 世紀 COE プログラムが始まり、大型研究資金等を得ることによって、慶應義塾大学、一橋大学、大阪大学、東京大学、お茶の水女子大学等多くの大学がパネルデータ調査を実施し始めた。

同時に厚生労働省では「21 世紀縦断調査」の名のもとに、2001 年生まれの新生児、若年者、中高年者といくつかの年齢グループをターゲットにパネルデータ調査を始めた。また、経済産業省と一橋大学が共同で Japanese Study of Aging and Retirement (JSTAR)(通称「くらしと健康の調査」)を 2005 年より開始し、50 歳以上 75 歳未満の高齢者 4200 名の健康状態や退職行動を調査している。

これらの調査では、パネル調査先進国の欧米での経験を参考に、調査設計、回収率の確保、サンプル脱落問題などに取り組んできた。多くの問題については想定できたが、中には、想定外の問題もあった。本稿では、そのような日本でのパネルデータ調査の経験から得られた知見を紹介し、皆さんと情報を共有したい。

言うまでもなく、私に関わった調査は多くのパネル調査の一部であり、必ずしも全体像を示すものではないことを、あらかじめお断りしておきたい。

2. パネルデータの調査設計

パネルデータ調査も基本的には統計調査法に基づいて設計され、実施されている。具体的な内容は調査によって違ってくるが、一般的な流れとしては次のようになっている。

(i) 調査目標の設定と調査票の作成

調査の目的と具体的な調査内容が設定され、それに応じて誤解なく答えられるように、後で解釈に困るようなことのないように、適切な質問項目を作る。その際、平均的な回答時間を考慮して質問の数や配列を決め、また個人属性に関する質問にも十分配慮することが必要である。調査票

には調査対象者の属性に関する質問(フェイス項目と呼ぶ)が含まれる¹。調査の内容に関わる項目は、回答内容を予め選択肢として用意したプリコード項目と質問に関する回答を自由に記入させる自由記述項目に大別できる。プリコード項目は回答しやすい反面、回答の選択肢が限定されており、予想外の答えには対応できないという問題がある。具体的には、多肢選択法(一つだけ選択)、無制限複数選択法(複数選択可)、制限複数選択法(選択数を制限)、完全順位法(全てに順位を付ける)、一部順位法(一部選択して順位付けする)などがある。自由記述項目は数値記入法と文字記入法がある。統計分析では数値記入法によるデータが最も用いられるが、質的情報としてプリコード項目の選択肢を用いることもある。質問の設定の仕方には気を付けるべき点もあるし、回答拒否を誘うような質問は回避し、事実と評価を区別することも重要である。

(ii)調査仕様の決定

調査の対象となる母集団を決め、抽出単位が個人か世帯か法人かを定める。調査対象を母集団全てとする場合には全数調査(悉皆調査)、一部を取り出す場合を標本調査という。全数調査には多大な費用と労力がかかるので、一般に標本調査が行われる。また調査期間や費用、調査対象となる標本数も決める必要がある。パネルデータを収集するためにはそれなりのノウハウが必要であり、実際には経験を積んだ各省庁統計担当部局あるいは民間調査会社が仕様を決めて、調査の予算に応じて標本数が決まっているようである。

(iii)標本の抽出

標本抽出の方法は無作為抽出法(random sampling)と有為抽出法(purposive selection)の2つがあり、一般には無作為抽出法が用いられることが多い。母集団の中から標本として等確率で抽出されるような方法が無作為抽出法とよぶ。この無作為抽出の方法はさらに分類することができるが、ここでは代表的な層別抽出法(stratified sampling)と2段階抽出法(two-stage sampling)について見ておきたい。層別抽出法は属性の構成比率の予備知識を利用して母集団を層別し、各層に対して、住民基本台帳などの台帳(フレーム)から乱数を用いて標本を抽出する。層別化しない単純無作為抽出法より精度が高いとされている²。2段階抽出法は母集団を地域によって1次抽出単位(都道府県、市町村等)に分け、まず1次抽出単位を抽出した後で、その単位から標本を抽出する。この2つの抽出方法を組み合わせた層別2段階抽出法では母集団をいくつかの層に分け、層ごとに2段階抽出を行うものであり、全国規模の主要な調査がこの方法を採用している。母集団から標本を抽出するためには、母集団全ての対象が含まれている台帳が必要になるが、世帯であれば、国勢調査や住民基本台帳に基づくことが多い。企業であれば、事業所統計調査、工業統計調査、商業統計調査などの企業センサスや国税庁資料に基づく営利法人名簿、帝国データバン

¹属性を表わす代表的なフェイス項目としては、性別、年齢、学歴、職業、年収、労働経験年数、婚姻形態、住居形態、住所、同居家族、兄弟姉妹の有無、職歴、父母の状況、子供の数などがある。

²JSTARでは、この方法を用いて全国から5市を選んで、その中で無作為抽出を行ってサンプルを選択している。この5市の選択にあたっては調査協力が自治体から得られることなどを理由としており、実際にサンプルの回収率が60%に達するなど、通常のパネルデータ調査の回収率をはるかに上回る結果を得ている。

クのデータベース等に基づいて抽出されることが多い。

(iv) 調査票による調査の実施

調査票に答えてもらうためには一般に次のような方法がとられることが多い。①面接調査 この調査は調査員が直接調査対象者のところまで出向き、調査への協力を依頼する。この方法では回答がその場で得られるため確実に情報が取れるし、調査員が記入するので間違いも少ない。しかし、限られた時間で質問するために質問項目は限定され、記憶に頼るような質問に対しては誤差が入ってくる可能性が高い。②郵送調査 この調査では調査票を郵送で送りつけ一定期間後に郵送によって回収するものであるが、対象者の時間の都合に応じて答えられるし、その分時間も多少かかるような質問もできる。しかし、対象者の自主性に依存しているために回収率は一般に低い。③電話調査 この方法は対象地域の電話帳から無作為に電話番号を抽出して対象者を選ぶものであり、調査費用は上の2つと比べるとかなり安くつく。しかし、電話帳から無作為抽出したとしても、対象地域に住む住民が特定の属性や時間帯によって回答者が偏る場合には問題となる。④留置調査 調査員が調査対象者の住居を訪問し、調査の主旨を説明し、調査票を配布し一定期間内に記入しておくことを頼む。一定期間後、調査員が調査票を受け取りに再び訪問し、調査票に記入漏れがないかどうかを確かめた上で、調査票を回収する。この方法は面接法に近いが一定期間時間を置くので、かなり多項目の調査も可能になる。また、調査員が複数回訪問することで、信頼関係が生まれることも調査にとってはメリットである。パネルデータ調査のように継続的に同一対象者に調査を依頼する場合には、費用はかかるが、この留置法によることが望ましい。⑤集合調査 これは一定の場所に対象者を集めて回答してもらう方式で、学校、会社、病院など人が集まってくる場所が対象になることが多い。これも標本の代表性という意味では問題があるが、一度に回答が回収できるということで費用節約的な調査方法である。

(v) 調査結果の編集・集計

回収された調査票には様々な誤差が入っており、データとして入力する前にそれらの誤差を出来る限り修正する。単純な記入ミスや回答方法の誤解などで適切な回答が類推できる場合には修正を施す。質問とは関係のない答え方をしているものなどは無回答扱いとする。これらの作業は個々の問題に当たってみなければ一般的な解決方法が在るわけではない。調査実施者は修正に恣意性が入らないように厳密な手続きを決めておく必要がある。有効な調査票を標本数で割り、回収率を計算しておくことも重要である。またパネルデータでは累積した脱落者数や脱落率も把握しておくべきである。

パネルデータは具体的には次の3種類の調査方法によって集められている。

(i) クロスセクション調査で調査対象が複数回の調査で重複しているケース

このタイプの調査はパネルデータを作成する目的で行われたわけではなく、一定の条件を満たす経済主体が必ず調査対象となるようにデザインされたものである。例えば、証券取引所に上場

しているすべての企業は『有価証券報告書』を財務省に提出する必要があるが、『有価証券報告書』の企業財務データを同一企業について複数年つなぎ合わせれば企業のパネルデータを作ることができる。同様に経済産業省で調査している『企業活動基本調査』は資本金 2000 万円、従業員数 50 名以上の製造業を中心とするすべての企業を調査対象としている。この調査でも、従業員数が 50 名以下になるか、廃業するのでなければ必ず繰り返し調査対象になるので、事後的にパネルデータとして再構成することが可能になる。これらの調査では一定の条件を満たす主体がすべて調査の対象になるという意味では標本調査ではなく(ある種の全数調査・悉皆調査)、調査対象が途中で理由もなく脱落するという問題もほとんどない。しかし、各年の連続性は意識されておらず、回答者も年によって交代することも多いので、場合によっては、回答者の理解の違いや記入誤差によってデータが大きくぶれることもある。このタイプのデータを利用する場合には、データの非連続性が本当の変化なのか誤差なのかを注意深く吟味する必要がある。

(ii) クロスセクション調査で調査対象が一定期間継続して調査に参加し、一定の割合で調査対象が入れ替わるケース

この調査はパネルデータ調査と考えることもできるが、一般的にはクロスセクション調査として設計されている。具体的な例としては総務省の『家計調査』がある。この調査では 6 ヶ月間同一の家計が家計簿をつけ、毎月 6 分 1 のサンプルが入れ替わる。詳細な家計簿を 6 ヶ月連続して付けることには、調査対象にかなりの負担を強いることになるが、海外の同様の家計調査ではインタビュー形式ではるかに短い期間(例えば 1 週間分)の消費について調査しているのに過ぎないことと比べると信頼のおける調査となっている。また毎月の調査の連続性という意味でも 6 分の 5 が前月と同じ家計であることから、標本の交替による不連続性は小さい。この調査は全国の二人以上の一般世帯および単身の一般世帯を対象に、全国の市町村から調査市町村を抽出し、抽出された調査市町村から調査単位区を抽出した後に、調査単位区の中から調査世帯を抽出する層化 3 段抽出法によっている。調査世帯の交替は 1 調査単位区 6 世帯を単位として全国で毎月 6 分の 1 ずつ行っている。

(iii) パネルデータ調査で調査世帯の交替は行わない

これは当初より同一主体を継続的に調査し、統計を蓄積することを目的に設計されている。このタイプの調査としては、例えば、財団法人家計経済研究所の『消費生活に関するパネル調査』(通称 家計研パネル)がある。この調査は 1993 年から毎年実施されており、1993 年時点で満 24 歳から 34 歳の 1500 人の女性をコーホート A として、1998 年時点で満 24 歳から 27 歳の 500 人の女性をコーホート B として追跡調査している。調査は全国の都道府県を 8 ブロックに分類し、そのブロックを都市規模によって 13 大都市、その他の市、町村に分類した層化 2 段無作為抽出を行っている。また調査票を留め置いて一定期間後に回収するという留置法を用いているのでインタビュー形式に見られる記憶違いなどの問題は少ない。しかし逆に時間が経つにつれてサンプルが脱落していく問題はある。この点については後述する。

調査上の問題点

既に何度も論じてきたように、パネルデータ調査は同じ対象を繰り返し調査するという点でこの対象の時間を通じた変化を捉えることができ経済行動を分析する上で非常に有益ではあるが、そのことは同時に時間を通して回答者集団の母集団に対する代表性が次第に失われていくという問題を抱えていることも意味している。

この問題はいくつかの理由で生じる。第一にパネルデータ調査に対して慣れてくることによって回答に歪みが生じる可能性がある。第二に以前の調査の回答に縛られて正直な回答ができないということも考えられる。第三に次第に調査に参加することがわずらわしくなり回答拒否(脱落)するようになり、第四に転居などによって追跡が難しくなるということも考えられる。

パネルデータ調査において代表性が確保されているかどうかは5つのレベルで検討されるべきである。

(i) 標本設定時脱落による歪み

パネルデータ調査では母集団から無作為抽出した標本に対して、調査に先立ってモニターの受諾を確認する。この時点で拒否されるケースを標本設定時脱落という。この結果、脱落した標本が以後のパネルデータ調査にもたらず歪み(偏り)を測定することは難しい。というのは脱落した標本からは一度も調査を行っていないからである。しかし抽出過程で性別、年齢、地域などの住民代表ベースの情報が用いられていれば、それを用いて調査不能になったグループと調査回収できたグループを比較し、調査不能グループに対しても調査に対する回答を予測(推定)することが可能になる。この推定結果と実際のパネルデータ調査の結果を対照すれば、調査不能グループのもたらした歪み(偏り)が推計できる。このように、標本設定時に脱落したり、調査のかなり初期に脱落するグループに対しては母集団の同位置層から代替標本を無作為抽出して補填することが多いが、その新たに選んだグループが調査不能グループの歪みを補正していることを確かめることが必要になる。

(ii) 継続時脱落による歪み

ある程度、調査を継続したのちに何らかの理由で脱落する標本もある。これがまったくランダムに発生しているのであれば大きな問題ではないが、脱落が一定の理由によるシステムティックなものであれば、それは問題を含んでいる。この歪み(偏り)を評価するためには、途中で脱落した調査継続標本が脱落する前までに回答していた数値を調査継続して脱落してない標本と比較することで、その歪みを評価することができる。

(iii) 調査慣れがもたらす歪み

調査慣れや以前の調査の影響によって回答にどの程度歪みがもたらされているかは、新たに

無作為抽出した標本と比較することで評価できる。この評価のために新たな標本を導入するということは費用もかかるし、実際の手間も大きい。この種の歪みが大きいとわかっている場合には、調査自体に慣れを生じさせないような工夫、過去の調査の影響を少なくする質問の仕方を考えるべきであろう。

(iv) 回答者の同一性の確認

パネルデータ調査を訪問留置法によって行う場合、標本抽出された本人ではない他の家族が回答するケースも見られる。具体的な数値データであれば、矛盾に気づくことも多いが、意識調査に対する回答に別人の回答がパネルデータとして入ってくると、深刻な誤差を生じさせることになり、この問題に対しては回答者が本人であるかどうかの確認を調査票に入れることが重要である。

(v) 回答誤差

パネルデータ調査で同一の質問を複数回に亘って行う場合、回答に誤差が見られることがある。もちろん本当に意見が変わる場合もあるだろうが、回答者が違っていたり、回答時点での心理状態が違うといったことも考えられる。パネルデータ調査における各回のマージナル分布が同じで、前回と今回のクロス表がほぼ対称という条件を満たすならば、各回答者には本来の態度があり、態度の強度に応じていくつかの集団に分類され、態度強度が異なると質問に対する回答選択の確率が異なる、各質問に対する回答誤差は独立である等を仮定して回答誤差を推計するモデルを構築することができる。

3. 脱落問題

上述の調査上の問題点の中でも、標本が一定期間後に脱落していくケースは広範に見らるが、この問題が検討されることは、これまで統計実務家など一部の関係者に限定されていた。しかし、近年、パネルデータの利用が増えるに従い、またパネルデータの蓄積が進むに従い、脱落サンプルの問題は認知されるようになってきた。

脱落サンプルの実態

まず限定的ではあるが、脱落サンプルが実際のパネルデータ調査でどの程度起こっているのかを確認しておこう。パネルデータ調査の先進国であるアメリカでは代表的なパネルデータ調査である The Panel Study of Income Dynamics (PSID) に関して脱落サンプルの問題が詳細に検討されている、1968 年には 4802 家族が標本抽出され³、翌年には 88%の家族が残り、12%が脱落している。以後、1989 年に至るまで年率 2.5-3%が脱落し、1989 年時点で 49%の家族が残り、51%

³ 4802 家族の内、2930 家族(5 分の 3)がミシガン大学の The Institute of Social Research 附属の Survey Research Center (SRC)の台帳(フレーム)から選択され、残りの 1872 家族(5 分の 2)が The Bureau of the Census の Survey of Economic Opportunity(SEO)に含まれている低所得家計(SMSA に入っている)台帳から選ばれた。

が累積して脱落していった。1993年、25年後時点で残っているオリジナルサンプルは約10%であると言われている。すなわち、四半世紀後にはオリジナルサンプルの90%はいずれかの時点で脱落してしまったことになる。

脱落の理由は転居、死亡、家族全体の非回答などが挙げられている。脱落者の属性を分類すると、(1)社会扶助、(2)未婚者、(3)高齢者、(4)有色人種、(5)低学歴、(6)労働時間が短い、(7)低賃金、(8)借家住まい、を満たす家族である可能性が高いことがわかった。これは、脱落者が一般的には社会的に低い地位にいる可能性が高いことを意味している。しかし、脱落者の中には高所得者も含まれていることも確認されている。

我が国のパイオニア的パネルデータ調査である『消費生活に関するパネルデータ調査』(通称家計研パネル)(家計経済研究所)でも脱落サンプル問題が出ており、平成15年度版『家計・仕事・暮らしと女性の現在：消費生活に関するパネルデータ調査(第10年度)』の第III部で、この調査がどの程度問題となっているかを論じている。この調査は1993年より始まっており、2003年で10回の調査が行われているが、その内訳は、1993年時点で24-34歳1500人で始まったコーホートAと1998年度から24-27歳500人で始まったコーホートBに分かれる。その内、脱落サンプルはコーホートAで471人、コーホートBで175人、合計656人となっている。脱落比率に直すとコーホートAが31.4%、コーホートBが35%となっている。この数字をアメリカのPSIDの10年目の数字である30.3%と比べるとほぼ同程度の脱落率であることがわかる。

村上(2003)は脱落理由を分析している。過半数を占める「詳細不明」を別にすると、コーホートAでは「多忙」(25.1%)、「転居先不明」(12.5%)、「(長期)不在」(10.6%)となり、コーホートBでは「多忙」(24.1%)、「転居先不明」「(長期)不在」「結婚」がそれぞれ12.0%となっている。さらにデータを詳しく見ると、調査の初期の段階では「家族の反対」が多く、時間がたつにつれて就職・転居・結婚などのライフイベントの発生によって脱落していく傾向がある。復活したサンプルと復活しなかったサンプルを比べると、「死亡」「転居先不明」などのように物理的に不可能な場合、調査そのものに不信感、負担感がある場合は復活していない。「結婚」「離婚・別居」「転居」「家族の反対」などの理由を挙げる人も復活していない。「結婚」「家族の反対」は結婚相手が調査に反対するケースが多いと言われている。これは質問が本人のみならず、結婚相手やその両親にまで及ぶためであると思われる⁴。「離婚・別居」「転居」を機に脱落するのはいろいろな意味で心機一転したいので、継続的な調査もやめてしまいたいということであろう。逆に、復活しうるのは「(長期)不在」「入院中」「体調不良・病気」「多忙」「出産」「就職・転職」「家族の病気・不幸」など脱落理由が一時的なものである場合に限られる。

坂本(2003)は脱落サンプルの統計的特性を考察している。方法としては「継続回答者」と「脱落

⁴日本の場合、広い意味で結婚を理由に脱落しているサンプルが無視できないほどある。結婚前後の労働供給や出産、育児などの行動を分析することは、女性パネルデータ分析では焦点となるトピックであり、それに該当するサンプルが脱落していくことは統計調査としても重要な問題である。村上(2003)が指摘しているように、対象者およびその家族に調査の意義を理解してもらい信頼関係を築くこと、対象者が多忙な場合にも回答が得られるような工夫をすること、対象者に感謝の気持ちを伝えることなど、地道な努力が必要であることが強く認識されている。

者」を分け、前年の回答データの平均値を比較している。それによると、脱落しやすい属性は、(1)年齢が若い、(2)未婚者、(3)低学歴、(4)有業者、(5)高所得者、(6)子供の人数が少ない、(7)結婚予定者、(8)核家族、となっている。これらの理由は一部アメリカの脱落理由と重なるが、有業で高所得者ほど脱落しやすいというのは逆である。また結婚を機に調査から脱落するというのもこの調査の特徴となっている⁵。

このように、パネルデータからの脱落という同一の行動を見ても、それぞれの国、それぞれの調査によって脱落理由もパターンも異なっている。しかし、脱落率の低い調査はいずれも調査機関が調査対象に対して積極的にアプローチし、移動しても追跡調査するなど、かなりの努力をしていることが見て取れる。これはパネルデータ調査を継続的に行う上で重要な点であり、この眼に見えない努力がパネルデータの利用価値を高めていることを認識しておく必要があるだろう。

脱落問題に関連して、東日本大震災の被災者あるいは本人は被災していなくとも被災地域に居住している人の反応はどうかという点について情報を提供しておきたい。

『慶応家計パネル調査(Keio Household Panel Survey: KHPS)』は2004年1月より調査を開始し、2011年3月で8回目の調査を終了したところである。2011年3月11日の東日本大震災では、被災地域にも調査対象サンプルが含まれており、回答の回収に時間がかかったが、被災地域以外では前年度からの回収率が94.0%(2968サンプル中2789サンプルが回答)に対して、被災地域では91.3%(264サンプル中241サンプルが回答)となり、若干の落ち込みは見られるが、大きな差ではないと調査を行った中央調査社では判断している。実際には、多くの調査は3月11日以前に終了しており、被災によって回答不可能になったというより、調査員が被災したことで調査票の回収が難しくなったケースが見られるということである。

家計経済研究所の『消費生活に関するパネル調査』は毎年9月に調査を行っており、慶応パネル調査のように直接的に調査期間中に大震災の影響を受けた訳ではないが、逆に大震災後時間が経っており、居住地域の変更や離職などが発生している可能性があり、8月現在では所在確認の意味も含めてお見舞い状を郵送しているところである。

もう一つ論じておきたいのは、脱落率によるが、10年ぐらいで半減するとすれば、サンプル数がある程度確保したければ10年に一回ぐらい新しい調査サンプルを追加する(new waveの追加)が必要が出てくる。家計研パネルではすでに3回(第5回(1997)、第11回(2003)、第16回(2008))新しいウェーブを追加している。慶応家計パネルでも2007年度に1400件の新しいサンプルを追加している。また、21世紀縦断調査でも2010年5月10-24日(平成22年出生児)に生まれた全ての新生児を新ウェーブとして追加している。

このように、途中で新しいウェーブを追加してサンプル数を確保するという考え方は正当化できるのだろうか。もちろん理想的には、最初にサンプリングされた対象が脱落せずに継続的に回答してくれれば、パネルデータとしての蓄積は進み、経済主体のライフサイクルを通じた変化をより

⁵ 同様の脱落問題に関する研究は慶応家計パネル調査でも行われている。

厳密に追うことが可能になる。しかし、実際には脱落という現実があり、サンプル数が低減していく事実は何の対処もしないですまされるものでもない。そのような事実に対する対処法として新しいウェーブの追加が行われているのである。

パネルデータのサンプルサイズ維持の難しさが、日本でパネル調査を実際に行ってみて次第に実感としてわかってきたのであるが、そもそもパネルデータとして何年ぐらいのデータ蓄積が必要なのか(期間Tのサイズ)についても考える必要がある。例えば、クロスカントリーのパネルデータ、企業財務のパネルデータであれば、景気循環の1サイクル程度(5-8年)あれば、何らかの分析が出来るのに対して、個人ベースのパネルデータであればライフサイクルを通してのデータ蓄積が望ましい。しかし、次節で述べるように調査研究費は無限に続くものではない。これに対する解決策は第5節で論じたいと思う。

4. 政府統計と大学を中心とした研究機関の統計との違い

政府統計は統計調査自体に厳密に目的が設定されており、それから逸脱した質問は出来ない。政府省庁間の縦割り行政のために、個人や家計を総合的に調査することが難しい。それに対して大学の統計では、調査目的に強い縛りがなく、総合的に調査することが可能である。しかし、政府統計ほどの行政上の強制力はなく、回答拒否率も高いなどの弱点もある。

大学で集めているパネルデータは大型科学研究費等を財源として収集されているものであり、大型科研の支給が打ち切られると維持することが難しくなる。近年の大型科研の資金配分の考え方は、拠点となる大学間での競争をベースにしており、それぞれの拠点が特色のある研究テーマを提示し、それに応じたパネルデータ調査を行っている。従って、どこかの拠点での調査が維持できなくなっても、他の拠点での調査で取って代わるということは出来ないというのが実情である。

各大学は各自で開始したパネルデータ調査を今後とも維持するか、ある程度で区切りがついたところで終了するかの決断を迫られている。また継続を決断した大学は資金調達と後継者の育成に努めなければならない。かなりの規模の調査費を競争的資金として、獲得し続けるのは相当難しいと言わなければならないだろう。大学を中心としたパネルデータ調査は危うい基盤の上に乗っているということを知っていただきたい。

もちろん政府統計も財政削減のあおりをうけて大幅な見直しが行われており、こちらでも明るい見通しは立っていない。指定統計が突然打ち切りになるというような状況は何としても回避する必要があるし、承認統計であっても、「21世紀縦断調査」のように、政府が作っている唯一の家計パネルデータであり、このデータを維持することの意義は極めて大きい。

政府統計の重要性は広く認識されているし、大学統計と比較してもその規模も内容も充実しているが、実際の有用性はその利用によって証明されるべきものである。政策立案に関連して決定的に重要な研究が政府統計をつかって行われたというような実績が積み重なって初めて、政府の政策立案にとってこれらの統計が不可欠であると論じることができるのではないだろうか。この点では、政府統計の公開、2次利用の促進が強く望まれるところである。

もう一つ、政府統計と大学統計の違いについて触れておきたい。それは、パネルデータの性格を反映した設問になっているかということである。具体的には、パネルデータの特徴は、①継続して同じ質問を繰り返し、②過去にさかのぼって経験を聞く、③政策変更に対する反応を聞くことができる場所にあり、その特徴を生かすことによって、これまでのクロスセクション統計にない分析が出来るのである。

ところが、厚生労働省が実施している「21世紀縦断調査」ではパネルデータとしての設問の意識が低いのではないかと危惧される程、継続して調査している変数が極めて少ない。実際に、『21世紀出生児縦断調査』では出生児の身長、体重、親の収入、子育て費用などが継続してとれる数少ない変数である。この統計は統計行政上は指定統計ではなく承認統計であり、実施機関内だけで調査項目が設定でき、外部から質問事項に関するチェックが必要のない統計になっており、パネルデータとしての価値が損なわれているように思われる。ただ、『21世紀出生児縦断調査』のために弁明しておく、通常のパネル調査は成熟した大人の家計行動の継続的な調査であり、大きな変化は雇用状況や健康状態を除けばそれほどないので、ある程度、同じ質問を繰り返すことが意味がある。それに対して、2001年の出生児に関しては、成長の記録として、身長体重などを記録する他には、年とともに変化していく子供の関心事や時間の使い方などを聞く必要があり、これは毎年変化していくものであり、必然的に質問内容も継続性がなくなることはやむを得ない面がある。また、この年齢からのパネルデータ調査は世界でも珍しい試みであり、必ずや長期的には価値のあるパネルデータとなると評価できる。

大学発のパネルデータ調査は、家計経済研究所の『消費生活に関するパネルデータ調査』での経験もあり、パネルデータとしての特色を生かしたものになっているように感じられる。また、設問自体に大学院生を含む大学研究者が直接に参加できる機会があるということは、きわめて貴重な経験になっており、広い意味での後継者育成の役割を果たしていることも重要である。

5. 今後のパネル統計調査のありかた

政府財政がひっ迫しており、潤沢な資金が統計調査に回るような環境にないことは周知の通りである。また、少子化、核家族化、単身世帯化する中で、統計調査を従来の方法で行っている回収率や脱落問題から解放されることはない。

これらの問題に対して、完全な解決策は無いが、個人識別番号(通称「マイナンバー」)、企業識別番号を導入してレジスター情報(行政記録情報あるいは業務情報)からパネルデータを作る仕組みは魅力的な代替策である。

Røed and Raaum (2003)が論じているように、北欧では個人識別番号をマッチング・キーにして、様々な組み合わせのパネルデータを作ることが可能になっており、実際に作られている。家計関連の情報のうち、所得、社会保障、税、教育、労働(雇用)関係のものはレジスター情報としてかなり手に入るの、基本的に個人識別番号を持つすべての市民に対して、パネルデータを作成することができ、これまでのように新規のパネルデータ調査を費用をかけて行う必要はなくなる。また、

行政記録であれば、脱落や回答拒否は原則としてない⁶。

一般的な行政記録情報や業務統計をパネル化するためには、そのデータが長期間にわたって保存されていることが必要だが、多くの業務統計は、これまで情報保存コストがそれなりに高かったために、一定期間後廃棄されることが多く、また他の目的で利用することを考慮していないがために、加工して目的外利用することが極めて煩雑になっている情報が多い。

すなわち、これまで行政記録・業務統計は各省庁が独自の業務目的のために、データ管理会社を選択し、全く他の利用を考慮に入れない汎用性のないデータベースを構築し、利用してきた経緯がある。レジスター情報を利用するためには、この業務目的に特化したデータベースを汎用性・互換性・接続性のあるものに変えていく必要がある。つまり、レジスター情報を別目的で利用することを視野に入れて、一般目的技術 (general purpose technology) を用いたデータベースへの変換が求められているのである。

また、同時に年金記録問題で明らかになったように、業務情報を当事者がチェックできる仕組みを作る必要がある。これによって行政記録の自己点検、修正が可能になる。

レジスター情報を用いてパネルデータを作ることは、識別番号を導入すれば即解決できる問題ではなく、さらに広範な省庁間のデータベースの調整が必要になることは重ねて指摘しておきたい。しかし、これらのコストをかけても長期的にレジスター情報からパネルデータを作成するという試みは安価で確実な統計情報の収集方法であると考えられる。我が国政府には一考を促したい。

6. データアーカイブの設置

政府統計であれ大学統計であれ、莫大な資金を投入して収集されたマイクロデータ、パネルデータの維持管理のためのデータアーカイブセンターの設立が望まれることは言うまでもないだろう。

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、戦前期からの政府統計資料およびその関連の地域データ、マイクロデータを収集し公開している。しかし、政府統計のデジタル情報を継続的に保管し、維持管理するという機能は、残念ながら今のところ果たしていない。多くのデジタル情報も、時間がたつにつれて、適切に維持保管しなければ劣化して、使用に耐えなくなる。政府統計はある程度の時間がたつと顧みられることがなくなり、維持管理がおろそかになる。このようなデータを維持管理するには大学の研究施設が最適だと思われる。

東京大学社会科学研究所附属社会調査・データアーカイブ研究センターでは、民間研究機関や政府各省が独自に行った統計法に係らない調査の1次データをSSJデータアーカイブとして収集し、保管、公開してきた。

これらの大学機関の実績とノウハウを生かしながら、パネルデータの保管、維持、公開が責任のある機関で行われることが望ましいだろう。

例えばパネルデータ収集が終了し、研究プロジェクトが解散したとしても、集められたパネルデータ

⁶ レジスター情報の統計データへの変換に関しては Wallgren and Wallgren(2007)を参照。もちろん、すべての個人行動がレジスター情報として記録されている訳ではないので、消費等の個人行動の把握については標本調査が必要となる。

は後の世代の研究者が、全く別の問題意識から利用したいという要請が出てくる可能性がある。とりわけ1990年代から2000年代にかけては日本社会は劇的な変化を遂げた時期であり、この時期に日本社会で起こったことを統計的に解析する意義はかなり大きいものがある。それらの情報を伝える第1次資料であるパネルデータを保存しておくことは真剣に検討すべき課題である。

7. おわりに

日本における社会経済に関する実証研究が盛んになるにつれて、パネルデータを利用した研究が増えていくことは想像に難くない。政府統計も大学を中心とした民間統計も、その公開や利用を促進するという点では認識を共にしている。

1990年代にたまたま、パネルデータ分析という手法の面白さに引き込まれて研究を始めた者として、日本においてパネルデータがこれだけ蓄積されてきたことに、そして実際に多くのパネルデータ調査に関与してきたことに対して、個人的にも感慨深いものがある。また、その過程で、多くの新しい困難に直面してきたことはここで述べた通りである。

今後とも社会経済の問題が無くなることはなく、パネルデータを用いて解決策を立てようという試みは、その方法論や応用分野でも広がりを見せるだろう。そして、パネルデータの作成、質の改善、利用の促進、方法論上の向上などすべての側面において、多くの問題に直面することだろう。この点に関しては皆様のご協力、ご理解なしには一歩も進まない。絶大なるご支援と知的貢献を期待する次第である。

参考文献

- 北村行伸(2005)『パネルデータ分析』、岩波書店
- 坂本和靖(2003)「誰が脱落するのか—「消費生活に関するパネル調査」における脱落サンプルの分析」、財団法人家計経済研究所『消費生活に関するパネル調査(第10年度)―家計・仕事・暮らしと女性の現在』、国立出版局、pp.123-136.
- 村上あかね(2003)「なぜ脱落したのか—「消費生活に関するパネル調査」における脱落理由の分析」、財団法人家計経済研究所『消費生活に関するパネル調査(第10年度)―家計・仕事・暮らしと女性の現在』、国立出版局、pp.115-123.
- Røed, Kunt and Raaum, Oddbjørn (2003) “Administrative Registers-Unexplored Reservoirs of Scientific Knowledge?”, *The Economic Journal*, 113, F258-F281.
- Wallgren, Anders and Wallgren, Britt (2007) *Register-based Statistics: Administrative Data for Statistical Purposes*, Wiley.