

## エリア・サンプリングの問題点

氏 家 豊

Issues of Area Sampling

Yutaka UJIE

This article discusses two important problems of area sampling in public opinion surveys.

The first revolves around the weighting that should be done to obtain unbiased estimates by correcting for the unequal selection probabilities in households of different size. In practice, many Japanese surveys omit such weighting. When the response rate is low, analyses showed that weighted data had an effect on the proportions of other demographics.

The second revolves around the residential map itself. The currency of the residential map is extremely important because of its effect on the response rate and survey quality. Despite the relatively low population mobility in the area of the present survey, there were some changes due to persons moving in and out. Therefore, the residential map must be updated with current information prior to the fieldwork.

Key words: weighting procedures, degree of error, residential map, population mobility  
キーワード: ウェイト補正, ウェイト補正による誤差, 住宅地図, 転入転出

### 1. はじめに

「住民基本台帳法の一部を改正する法律」の成立(2006年)以来, 世論調査のための抽出台帳の利用が著しく制限された。そのため調査の実施においてエリア・サンプリングが採用される機会が多くなった。

エリア・サンプリングにおいては, 調査対象となる世帯を選び, その世帯の中から調査対象者を選ぶ。このとき, 世帯内に調査対象者となりうる人(以下, 対象資格者という)が何人いても, 調査対象者は1人だけ選ぶのが通常であるため, 母集団推計を行うための不偏推定値を得るべく, その世帯内の対象資格者の数によってウェイト補正を行う必要がある。ところが, 調査を実施する調査機関ではそのようなウェイト補正を行わないことがある。世帯内の対象資格者の数によるウェイト補正を行うことによって, 他の対象者属性の構成に偏りをもたらすことがあるためである。

しかし, エリア・サンプリングにおけるウェイト補正の実際のプロセスが明らかにされてこなかったために, 母集団の構成比からどれほどの偏りを生じるのかは不明であった。本稿では, このウェイト補正のプロセスを紹介し, ウェイト補正によって対象者属性の構成にどのように偏りが生じるのかを検証する。

また, そのようなウェイト補正の不都合は回収率の低下による標本構成の偏りに起因するため, 回収率の低下をもたらすいくつかの原因の中で基盤的な要因となっている抽出枠=住宅地図の鮮度について, その影響の有様を指摘する。

日本においてエリア・サンプリングは, 住民基本台帳の閲覧が制限される以前から, 経費の節減などの理由で用いられてきた。要は, 特定されたエリアから調査対象者を直接選出する方法であるが, 実際にフィールドワークを実施する調査機関によっても共通の抽出方法があるわけではない。おおよそは調査地点を定め, その調査地点の住宅地図を利用して調査対象世帯を選び, その世帯から1名の調査対象者を選ぶというのが一般的である。調査機関ごとに細部において様々なノウハウが蓄積されており, そのためかエリア・サンプリングに関する研究成果も多様な面を見せている。そ

こで今日までのエリア・サンプリングに関する先行研究をレビューすることによって、その方法と課題の多様性を確認することからはじめたい。

欧米をはじめ諸外国では日本のように住民基本台帳や選挙人名簿などの抽出台帳を利用した標本抽出が行えないところが多いので、面接調査などの調査員調査においてはエリア・サンプリングやクォータ法などがよく用いられている。したがって、それらのサンプリングの研究も古くから行われてきた。

Kish (1965) はその著“Survey sampling”において、サンプリングの基礎についてひととおり記した後で、サンプリング手法のひとつとしてエリア・サンプリングを紹介している。そこでは、調査の対象者が居住する住宅を抽出するための手続き全般について詳しく説明が行われている。住宅を抽出するためのエリアの設定から始まり、エリア内の区分によって、最終的に調査員が訪問する住宅を選出する手順と、調査員によるフィールドワークが混乱することなく実施されるための地図作成上の注意点とを具体的に説明している。Kish は、また、一世帯内に対象資格者が複数存在する場合に、その中で1人を調査対象者とするための選定方法についても、それまでデータを積み上げてきたミシガン大学の Survey Research Center の資料を元に検討を加えている。Kish らのこのような先駆的な研究は、日本におけるエリア・サンプリング理論の礎となっている。

日本におけるエリア・サンプリングに関する研究を概観すると以下のとおりである。

今から30年ほど前に、根本 (1983) はエリア・サンプリングの運用を説明しながら、世帯主年齢や家族人数などの調査結果について国勢調査との比較を行っている。ここでは、国勢調査区の範囲で設定された調査地点において「現地の住宅地図」を作成することによって調査対象者を抽出している<sup>1</sup>。その手順は、作成した現地住宅地図により調査対象世帯を選び、ランダムテーブルによる調査対象者個人の抽出を行うための導入質問である「リステイング用紙」を用いたものであった。調査時に調査対象者には「住宅地図を用いた調査対象者の選定」の仕方を説明し、調査に関する不安の解消に努めている。このように、調査対象者の名前がわからぬまま訪問することに対する調査員の不安を取

り除く工夫が考えられているが、導入質問までのアプローチの難しさも指摘されている。また、この調査以外に、八王子市の住宅地図から抽出した世帯と住民台帳との照合を行い、そのズレを発見したり、天理市の住民台帳から抽出した調査対象者を住宅地図上にプロットすることによって、そこでのズレを発見している。

JMRA (日本マーケティング・リサーチ協会) (2007) では、加盟機関のほとんどが市場調査を中心とした調査機関であることから、住民基本台帳法の改正以降、住民基本台帳や選挙人名簿を使用した調査がほとんど行えなくなったため、エリア・サンプリングの研究を熱心に行ってきた。その研究成果として「非名簿フレームによる世帯および個人の無作為抽出方法に関する国内外の先行研究事例をレビューすることにより、住民基本台帳に代わる抽出フレームの可能性を指摘する」ねらいで、「非名簿フレームによる無作為抽出法の国際研究」「日本国内における住宅地図フレームによる事例の紹介」「アメリカの事例の概観」「RDD による世帯内個人の無作為抽出法」「ビデオリサーチ社による研究事例の紹介」など幅広い報告が行われている。また、世帯抽出後、誕生日法で個人を抽出した調査 (調査の対象は首都圏40キロ圏の15歳以上75歳以下の男女個人、75地点において1,500サンプル) では、層別・県別に、調査地点における国勢調査の世帯数と住宅地図の世帯数を比較し、空き家や非該当など、住宅地図の不備による調査不能が少なからず発生していることを含めその差異を紹介している。さらに、調査対象世帯には依頼状のポスティングを行い、調査対象者個人の選定では誕生日法を用いたが、依頼状の出し方と誕生日法による個人の選定には調査実施上の課題が残り、工夫の余地があったと付している。この調査に関しては、調査不能についての報告も行われている。その報告によると、エリア・サンプリングにおいては、「拒否」や「不在」といった一般的な調査不能のほかに「住宅地図 (電子地図データベース) と現地の家屋の不一致」や「調査対象外 (80歳以上) のみの世帯」といった、住民台帳から抽出したときには起こり得ない調査不能があり、事前依頼が行えないために「オートロックで調査対象者本人に接触できず」「管理人・管理組合による拒否」といったものが含まれた。調査結果からは、住居形態が回収率に大きな影響を及ぼすことが明らかになり、集合住宅に対してどのような対策が立てられるかが、エリア・サンプリングにとって重要なポイント

<sup>1</sup> 調査地域は滋賀県ならびに奈良県、計40調査地点において580サンプルで実施。

であると指摘している。

エリア・サンプリングの実査上の問題について研究を続けているビデオリサーチでは、シミュレーションと実験調査からランダムテーブルを使用した個人抽出法が有効であることを示し(小柳, 2003), これに続けて、住民台帳を用いた抽出による従来型調査と、エリア・サンプリングにおいて世帯内の個人を選出する段階でランダムテーブルを用いた調査との比較によりエリア・サンプリングにおける問題点を指摘した。そこで調査の回収率を高めるためには、対象者が予め決まっていなかったときの「調査員のモチベーションアップ」のための方策や、効果的な訪問パターンを研究していく必要があることなどが確認されている(朝倉・小柳・染谷・桜井・武藤, 2004)。また、住民台帳による抽出、住宅地図とランダムテーブルを用いたエリア・サンプリングによる抽出、およびクォータ法の3通りの抽出方法による調査結果、さらにエリア・サンプリングによる調査において世帯の調査対象人数によるウェイト補正を行った結果、等について比較を行い、性・年齢、職業などの属性と「日常生活での行動・考え」や「所有商品」などに関する質問に対する回答の差異に関して、家族人数や職業といった実態については、住民台帳による抽出とエリア・サンプリングによる抽出でそれらの構成比に違いがみられた点を明らかにしている(朝倉・桜井・染谷, 2005)。

鄭(2002, 2007)は「世帯分布状況や地図に基づき個人サンプルを抽出する新たな確率抽出法を構築すること」を企図し、中国の北京市、上海市で実施した意識調査の事例を紹介している。そこでは、「抽出台帳が利用できない場合の個人サンプルの抽出方法を考案し、調査における非標本誤差について分析すること」を目的とし、あわせて「必要な回答標本数を確保するために、実際によく導入されている代替標本の性質を、調査データの解析によって明らかにする」ことを試みる。その後も東アジア諸国において新たな抽出方法を考案し実験調査を実施してきたが、それまでの知見に基づき、日本における標準的な標本抽出法と比較しながら新たなエリア・サンプリングの利用可能性を探っている。東京で実施した実験調査では、一方は住民台帳を用いた標準的な標本抽出法により、もう一方は住宅地図から調査対象世帯を選び、誕生日法により世帯内の個人を調査対象者として選出した<sup>2</sup>。サンプリングの操

作方法に工夫を加え、標本回収状況や誤差問題などの分析を行いながら、標本リストの整備、個人の抽出における調査員の恣意性の回避、調査員の訪問などに関する具体的な提言を行っている。

また、土屋他(土屋・前田・中村・坂本, 2007)が行った調査では、住宅地図により無作為抽出された住居の郵便受けに調査書類を直接投函するポスティングを用い、世帯内の個人の選定には任意法を採用して、エリア・サンプリングの方法上の可能性を追求している。この調査では、並行して同時期、同地域で同じ調査票を用いた郵送調査も実施され、両調査結果の比較も企図された。両調査とも、回収率の低い結果(ポスティング調査19%、郵送調査29%)となったが、ポスティングによる調査の実施方法における特徴や問題点が明らかにされ、今後のエリア・サンプリングの研究範囲を拡げた。

いうまでもなくエリア・サンプリングにおいては、調査対象者の抽出プロセスと調査世帯・調査対象者への協力依頼の方法が要となる。これまでの実験調査や研究はその点について多くの知見を提示してきた。特に、先行研究がそれまでの調査との継続性を見据えて、現地における抽出サンプルと従来の抽出枠である住民基本台帳との差異に留意している点は評価されよう。

このように多彩な研究が行われながらも、さらに検討すべき問題がいくつかある。世帯から選ばれる個人の抽出率が違うために行われるウェイト補正に関しての研究は、特に実際の調査データを用いたものとしてはあまり扱われていない。そもそもウェイト補正については、世帯抽出であるから電話調査と同じく「確率調整」をするのは当然であるとする考え(松田, 2008)が示される一方、実際の調査報告の中にはウェイト補正をしない集計結果が多く、それはウェイト補正が却って誤差を拡大する可能性が大きいからといわれている。そこで本稿では、エリア・サンプリングによる調査データを用いて、実際のウェイト補正の工程をたどり、そこにおける問題について検討する。

ところで、公表されているエリア・サンプリングによる調査データは非常に限られており、さらにウェイト補正に必要な世帯情報も完備されているとは限らず、当初の目論見どおりに分析を進めることはできなかった。それでも、エリア・サンプリングにおける個人の抽出率の違いによるウェイト補正において、何が、どのように問題になるのかを突き止めることができた。すなわち、個人の抽出率の違いから不偏推定値を

<sup>2</sup> 調査対象は東京都23区の20歳から79歳の男女個人、面接法による。

得るために行うべきウエイト補正の際に、ウエイト補正後の回収標本の属性で歪みが増幅する場合があることが確認された。

そこで、その原因となる回収標本の偏りを招く低回収率<sup>3</sup>の改善こそが、エリア・サンプリングにおける至急の課題であることを再確認し、そのための抽出手続上の基盤である住宅地図の問題に論及する。ここでは、時間的な経過により母集団を正確に反映した標本を得ることができない場合、回収率の低下をもたらす、回収標本の偏りを結果することになることを示した。

## 2. エリア・サンプリングの現状

最初に、世帯内で調査対象となる個人の選出方法については、調査実施機関によってそれぞれのノウハウがあり、一定の方式があるわけではない。

ここで扱うエリア・サンプリングは、調査実施方法として面接聴取法や留置法などの調査員調査を前提に、第1次抽出単位を調査地点、第2次抽出単位を世帯とし、各世帯から1人の調査対象者を抽出することを考えている。ここでは、その第2次抽出単位(世帯)の抽出のため、現在よく使われている住宅地図による方法について論じる。

第1次抽出単位(以下、PSU という)となる調査地点は、調査員の稼働に無理のない調査範囲が設けられることが必要のため、国勢調査区や投票区とすることが一般的である。国勢調査区も投票区も住所で範囲が指定された広がり(例えば「盛岡市中央通3丁目5~10番地」)である。マスコミ各社が行う調査では、選挙調査のために投票区単位とすることが多いが、国の世論調査の場合は国勢調査区をPSUとして用いる。国勢調査区は、1つの調査区が平均50世帯程度で全国に約98万ほどある。エリア・サンプリングにおける調査地点(PSU)の抽出については、国の世論調査における層化2段抽出法の抽出手続きと同様と考えてよい。すなわち、地域別、都市規模別に標本数を比例配分し、その標本数に見合う調査地点数を算出し、その数だけの国勢調査区を抽出する。

<sup>3</sup> 例えば、内閣府が2009年5月にエリア・サンプリングで実施した「地球温暖化対策の中期目標に関する世論調査」(全国20歳以上の者4,000人を対象に面接法で実施)では回収率が30.6%であった。また、前述のJMRAの調査(首都圏40キロ圏の15~75歳の者1,500人対象に留置法で実施)では16.5%であった。

抽出された調査地点において調査対象者を抽出するには、まず調査対象世帯を決め、その世帯内から1人の調査対象者を選出するが、このとき調査世帯の選出に際して住宅地図を用いることが多い。住宅地図を用いずに「エリア内で5軒おきに対象世帯を選べ」といった指示は、道路が入り組んでいたりすると調査員には大変な負担となる。住宅地図で予めマークされた対象世帯を探すことは調査員にとっての負担が軽減されるばかりでなく、調査員の恣意的な選定を排除した厳密さが確保されることになる。この住宅地図に記されている世帯がエリア・サンプリングにおける第2次抽出単位となる。従来は紙ベースの住宅地図にマークをして調査対象世帯を定めたが、現在は電子媒体の住宅地図を利用することができ、ここでは該当範囲の地図上に示された世帯を「世帯一覧」として列挙し、それらの世帯から起算番号と抽出間隔により調査対象世帯を抽出することができる。紙ベースの住宅地図上の世帯に起算番号と抽出間隔をあてはめて抽出するのと原理的には変わらないが、道路が複雑に入り組んでいたり住宅が散在したりしている場合は、紙ベースの地図上で起算番号や抽出間隔を当てはめる住宅の順番が抽出者によって一定でないことがあり得る。それを、整然とした世帯の一覧表として用意しておけば混乱がない。そして、その一覧表から抽出された世帯が一覧表とリンクしている住宅地図上の世帯にマークが入るようにしておけば、調査員には抽出世帯がマークされた住宅地図を提供することができる。

次に、対象世帯から1人の調査対象者を選出する。ここでは大別して2つの方法がとられている。ひとつは世帯内の対象資格者の数にかかわらず、無作為に個人が1人定まる指標を用いる方法である。この指標には誕生日を用いることが広く採用されている。例えば、対象となった世帯の誰かに「同居家族の20歳以上の人のうちで、今日以降一番早い誕生日の人は誰か」という具合に質問する。もうひとつの方法は、調査の対象資格者が何人いるかの情報を用いて無作為に1人の対象者を決める方法である。これは例えば、調査対象世帯に番号を振り、該当番号の世帯における対象資格者の数によって、世帯の中の何番目の人(年齢順)を調査対象者とするかをランダムテーブルにより定めるものである。

このようにして、調査員による恣意的な対象者選定の余地なく、調査対象者個人が選出される。したがって抽出台帳を用いたサンプリングと同様に設定標本数

が存在することになり、回収率が定義できることになる。

抽出台帳が利用できる場合と同様にリストアップされた調査対象者に調査するといっても、まったく同様なわけではない。抽出台帳によるリストアップの場合は、事前に調査対象者に調査の予定を知らせたり、前もって調査の協力依頼を行うことができるが、エリア・サンプリングの場合は指定された世帯を訪問して初めて調査対象者が特定されるので、事前に協力依頼を行うことは難しくなる。事前の協力依頼ができない場合の調査員の心理的負担や、回答者の側でも完全に不信感を払拭したうでの質問と回答ではないといった状況があり、このことによる測定誤差も考えられる。

なお、実際は調査機関など調査の実施者によって抽出手続が異なる場合があるが、本稿では世帯の抽出においては「住宅地図上で居住者のいる世帯だけを抽出の対象とする」という前提で話を進めることにする。

### 3. エリア・サンプリングの問題点

次に、ウエイト補正の問題点と住宅地図の正確さについてデータをもとに検討する。

#### 3.1 ウエイト補正の問題点

前述したように、世帯の中から1人の調査対象者を選出すると、その世帯に対象資格者が何人いるかによって個人の抽出率が違ってくる。世帯の規模によって選ばれる個人の抽出率に差があるので、抽出率が同じ場合に相当するように集計した結果が得られるようにウエイトで調整する必要がある。ただし、回収率が100%であればウエイト補正を行うことに問題はないが、通常の世論調査の場合はほとんどの場合100%の回収率ということはまずない。特に、世帯内対象資格者数による回収率の違いはそのままウエイト補正に直接に影響を与える。したがって、回収率が100%という前提が成立しないのでウエイト補正を行わないというケースが多いように思われる。

では、エリア・サンプリングの低回収率に基づく偏りを伴った回収標本においてウエイト補正を行うとどのような結果になるのか。ここでは、調査世帯の対象資格者の数によってウエイト補正を行った場合、調査結果がどのように変わるのかを補正の過程を通してみていくことにする。使用したデータは、新情報センターがアメリカのローパーセンターに供出しているデータ

の中のひとつで「SJC POLL 64」という世論調査<sup>4</sup>の結果である。この調査では、100地点のうち62地点において調査対象者を住民基本台帳から抽出しており、残りの38地点においてはエリア・サンプリングにより調査対象者を選んでいる。調査地点（PSU）は、1995年度の国勢調査結果をベースに調査年の住民基本台帳人口要覧を利用して母集団推計を行ったうえで、地域・都市規模別に層化<sup>5</sup>し、国勢調査区を抽出単位として無作為に抽出している。エリア・サンプリングにおける調査対象者は、これらの調査地点から選出された世帯の中で個人を指定する3段階抽出により抽出している。ただし、住民基本台帳による抽出を行った調査地点はその調査地点の市区町村で住民基本台帳の閲覧が許可されており、閲覧が許可されなかった市区町村の調査地点ではやむなくエリア・サンプリングを用いたのであり、2つの抽出方法を無作為に割り振ったのではないことを了解する必要がある。また、はじめからエリア・サンプリングにおけるウエイト補正を予定していないために世帯内の対象資格者数については情報がないので、質問の中にある「同居家族人数」をウエイト補正に用いている。したがって、本稿の目的であるウエイト補正の過程と数値の動きは明らかになるが、ウエイト補正值自体は単なる目安としておく。

ここでは、エリア・サンプリングによるデータだけを取り出し、そのデータについてウエイト補正を行った。実際の調査結果はウエイト補正を行っていないデータによる集計結果が発表されているが、ここではウエイト補正をしたら調査結果がどう変わるかをみた。また、ここで分析の対象とした調査結果は、母集団との比較が可能な性別、年齢、住居形態とした。

表1は、エリア・サンプリングの地点（以下、「修正前エリア地点」という）における性別の分布、および「修正前エリア地点」のデータについて世帯内の個人の抽出率の違いを家族人数で補正したもの（以下、これを「修正後エリア地点」という）の分布である。

「修正前エリア地点」と「修正後エリア地点」のそれぞれの調査結果について、性別の分布が母集団とどの

<sup>4</sup> 調査の対象は全国20歳以上1,500人、調査方法は留置法、1999年11月実施。回収数（率）は1,031（68.7%）。

<sup>5</sup> 地域は北海道から九州までの10区分、都市規模は東京都区部、東京都区部以外の政令指定都市、人口10万人以上の市、人口10万人未満の市、町村の5区分。

表1. 修正前後の「性別」分布と母集団における分布との適合度

	男	女	計	カイ2乗値
修正前エリア地点 (N)	193	209	402	0.03598
(%)	48.0	52.0	100.0	
修正後エリア地点 (N)	196	206	402	0.01204
(%)	48.8	51.2	100.0	

注) 自由度1のカイ2乗分布5%点 3.84146

表2. 「修正前エリア地点」における家族人数別「性別」

	男	女	計
1人世帯 (N <sub>1</sub> )	14	21	35
(%)	40.0	60.0	100.0
2人世帯 (N <sub>2</sub> )	48	47	95
(%)	50.5	49.5	100.0
3人世帯 (N <sub>3</sub> )	31	43	74
(%)	41.9	58.1	100.0
4人世帯 (N <sub>4</sub> )	53	49	102
(%)	52.0	48.0	100.0
5人世帯 (N <sub>5</sub> )	47	49	96
(%)	49.0	51.0	100.0
合計 (N)	193	209	402
(%)	48.0	52.0	100.0
ウェイト補正值 (N')	650	685	1,335
(%)	48.7	51.3	100.0

程度乖離しているのかをみるために適合度の検定 (信頼度 95%) を行った。その結果は表1に示したとおり、いずれにおいても有意差は認められなかった。

では、「修正前エリア地点」の「男」48.0%、「女」52.0%が家族人数によるウェイト補正によって「修正後エリア地点」の「男」48.8%、「女」51.2%に修正される過程をみてみることにする。表2は「修正前エリア地点」の性別分布を家族人数別にみたものである。これを用いて家族人数別の実数値 (N) を  $N_1 \times 1 + N_2 \times 2 + N_3 \times 3 + N_4 \times 4 + N_5 \times 5$  ( $N_i$  は家族人数  $i$  人の世帯からの調査回答者数) としてウェイト補正值を得るが、その値は「男」650、「女」685となる。これに対して、母集団の性別分布を世帯人員別<sup>6</sup>にみたものが表3である

表3. 母集団における世帯人員別「性別」

	男	女	計
1人世帯 (実数)	6,694,720	5,720,680	12,415,400
(%)	53.9	46.1	100.0
2人世帯 (実数)	10,811,610	12,175,830	22,987,440
(%)	47.0	53.0	100.0
3人世帯 (実数)	10,837,590	11,647,170	22,484,760
(%)	48.2	51.8	100.0
4人世帯 (実数)	10,599,970	11,359,830	21,959,800
(%)	48.3	51.7	100.0
5人世帯 (実数)	9,822,040	11,299,710	21,121,750
(%)	46.5	53.5	100.0
合計 (実数)	48,765,970	52,203,230	100,969,200
(%)	48.3	51.7	100.0

注記) 家族人数「5人世帯」は実際は5人以上世帯である。(以下同様)

が、これと対比させると表2の家族人数別にみた「性別」の分布が母集団における分布から偏っていることがわかる。このような偏りのまま家族人数によりウェイト補正された結果が表2の「ウェイト補正值」である。ところが、カイ2乗値をみてもわかるとおり、性別では結果的にウェイト補正が母集団の割合に近づいている。

次に、年齢についてはどうか。表4は「修正前エリア地点」と「修正後エリア地点」における年齢 (年代) の分布である。「40代」が「修正前エリア地点」(18.9%)から「修正後エリア地点」(22.3%)で3.4ポイントの増加を示し、「30代」でも「修正前エリア地点」(18.4%)から「修正後エリア地点」(20.8%)で2.4ポイントの増加となる。

年齢の分布(「20代」「30代」「40代」「50代」「60代」「70歳以上」の6区分)についての適合度検定結果は表4に示したとおりで、「修正前エリア地点」では有意差は認められず、「修正後エリア地点」では有意差が認められた。「修正前エリア地点」(カイ2乗値 8.97)と「修正後エリア地点」(同 23.11)の開きは大きい。

では、家族人数別の「年代」の区分を用いて、性別と同様にウェイト補正の過程をみてみよう。「修正前エリア地点」の家族人数別にみた「年代」が表5、母集団の世帯人員別にみた「年代」が表6である。表5をみ

る各人 (世帯員) を合わせた数<sup>6</sup>をいう。調査結果における (同居) 家族人数とはほぼ同義である。

<sup>6</sup> 国勢調査報告における世帯人員とは「世帯を構成す

表4. 修正前後の「年代」分布と母集団における分布との適合度

	20代	30代	40代	50代	60代	70歳以上	計	カイ2乗値
修正前エリア地点 (N)	77	74	76	66	66	43	402	8.97331
(%)	19.2	18.4	18.9	16.4	16.4	10.7	100.0	
修正後エリア地点 (N)	74	84	90	65	53	37	403	23.1138
(%)	18.4	20.8	22.3	16.1	13.2	9.2	100.0	

注) 自由度5のカイ2乗分布5%点 11.0705

表5. 「修正前エリア地点」における家族人数別「年代」

	20代	30代	40代	50代	60代	70歳以上	計
1人世帯 (N <sub>1</sub> )	12	6	2	4	5	6	35
(%)	34.3	17.1	5.7	11.4	14.3	17.1	100.0
2人世帯 (N <sub>2</sub> )	9	5	13	19	31	18	95
(%)	9.5	5.3	13.7	20.0	32.6	18.9	100.0
3人世帯 (N <sub>3</sub> )	22	11	6	13	16	6	74
(%)	29.7	14.9	8.1	17.6	21.6	8.1	100.0
4人世帯 (N <sub>4</sub> )	20	32	23	18	7	2	102
(%)	19.6	31.4	22.5	17.6	6.9	2.0	100.0
5人世帯 (N <sub>5</sub> )	14	20	32	12	7	11	96
(%)	14.6	20.8	33.3	12.5	7.3	11.5	100.0
合計 (N)	77	74	76	66	66	43	402
(%)	19.2	18.4	18.9	16.4	16.4	10.7	100.0
ウェイト補正值 (N')	246	277	298	213	178	123	1,335
(%)	18.4	20.7	22.3	16.0	13.3	9.2	100.0

表6. 母集団における世帯人員別「年代」

	20代	30代	40代	50代	60代	70歳以上	計
1人世帯 (実数)	3,651,260	1,866,420	1,341,160	1,747,260	1,592,380	2,216,930	12,415,410
(%)	29.4	15.0	10.8	14.1	12.8	17.9	100.0
2人世帯 (実数)	2,420,730	2,500,550	1,994,910	4,788,280	6,309,060	4,973,910	22,987,440
(%)	10.5	10.9	8.7	20.8	27.4	21.6	100.0
3人世帯 (実数)	4,241,820	4,102,940	3,182,140	5,425,840	3,378,170	2,153,870	22,484,780
(%)	18.9	18.2	14.2	24.1	15.0	9.6	100.0
4人世帯 (実数)	4,577,940	4,882,970	5,286,670	4,401,670	1,350,520	1,460,060	21,959,830
(%)	20.8	22.2	24.1	20.0	6.1	6.6	100.0
5人世帯 (実数)	3,299,980	3,553,500	4,947,800	2,846,600	2,274,390	4,199,540	21,121,810
(%)	15.6	16.8	23.4	13.5	10.8	19.9	100.0
合計 (実数)	18,191,730	16,906,380	16,752,680	19,209,650	14,904,520	15,004,310	100,969,270
(%)	18.0	16.7	16.6	19.0	14.8	14.9	100.0

表7. 修正前後の「住居形態」分布と母集団における分布との適合度

	一戸建	集合住宅	その他	計	カイ2乗値
修正前エリア地点 (N)	287	112	3	402	8.36776
(%)	71.4	27.9	0.7	100.0	
修正後エリア地点 (N)	307	92	2	401	16.5762
(%)	76.6	22.9	0.5	100.0	

注) 自由度2のカイ2乗分布5%点 5.99146

ると、「4人世帯」の年代区分で「30代」(31.4% ⇔ 母集団では22.2%)、ならびに「5人世帯」の年代区分で「40代」(33.3% ⇔ 母集団では23.4%)が母集団と比べて偏りが大きい。ウェイト補正により年代区分においても、このような偏りのまま家族人数によるウェイト補正値を得ることになる。その結果、「30代」「40代」では母集団の比率(それぞれ表6の「合計」に示されている16.7%, 16.6%)からますます離れていく。「70歳以上」では逆に「修正前エリア地点」(10.7%)から「修正後エリア地点」(9.2%)で減少となり、ここでも「30代」「40代」とは逆の現象で母集団の比率から離れていくことになる。

次に、住居形態についてみる。「修正前エリア地点」と「修正後エリア地点」の住居形態分布を示したものが表7である。「一戸建」が「修正前エリア地点」(71.4%)から「修正後エリア地点」(76.6%)で5ポイント以上の増加となる。

住居形態(「一戸建」「集合住宅」「その他」の3区分)についての適合度検定結果は表7に示したとおりで、ともに有意差が認められ「修正前エリア地点」(カイ2乗値8.37)と「修正後エリア地点」(同16.58)ではその開きも大きい。

「修正前エリア地点」の家族人数別にみた「住居形態」、母集団の世帯人員別にみた「住宅の建て方」の表は、それぞれ表8、表9である。母集団の分布と比べ、「修正前エリア地点」においては「3人世帯」および「4人世帯」で「一戸建」が6ポイント以上多くなってい

表8. 「修正前エリア地点」における家族人数別「住居形態」

	一戸建	集合住宅	その他	計
1人世帯 (N <sub>1</sub> )	11	23	1	35
(%)	31.4	65.7	2.9	100.0
2人世帯 (N <sub>2</sub> )	63	32	0	95
(%)	66.3	33.7	-	100.0
3人世帯 (N <sub>3</sub> )	53	20	1	74
(%)	71.6	27.0	1.4	100.0
4人世帯 (N <sub>4</sub> )	76	25	1	102
(%)	74.5	24.5	1.0	100.0
5人世帯 (N <sub>5</sub> )	84	12	0	96
(%)	87.5	12.5	-	100.0
合計 (N)	287	112	3	402
(%)	71.4	27.9	0.7	100.0
ウェイト補正値 (N')	1,020	307	8	1,335
(%)	76.4	23.0	0.6	100.0

る。このことが「修正後エリア地点」の「一戸建」の比率を生み出すことになる。

また、通常の世論調査においては「集合住宅」の回収率は往々にして低く、そのことにより、調査によっては対象資格者数自体の分布の偏りも大きくなる。そのようなときは、住居形態別の回収率の影響が誤差の拡大をさらに増幅することになる。

回収率が100%でしかも標本の属性に何ら偏りが無い場合には、世帯における個人の抽出率に違いがあるのだから、不偏推定値を得るために対象資格者数によるウェイト補正が必要である。ところが、回収率が低ければ対象資格者数(ここでは家族人数)の分布自体の偏りも大きくなることが考えられ、そのような状況でウェイト補正を行えば、他の属性の分布が母集団比率から一層離れるという現象が生じかねないことは

<sup>7</sup> 住居形態については、1999年の母集団推計値が不明なため、2000年の国勢調査結果(区分は「一戸建」「長屋建」「共同住宅」「その他」で、「長屋建」と「共同住宅」を合わせて「集合住宅」とした)を用いて検定を行った。また、世帯人員別性別・年齢・住宅の建て方の母集団データも同調査結果(13.8%抽出サンプルによるオーダーメイド集計のため、統計局公表の統計とは異なる)のものである。



表9. 母集団における世帯人員別にみた「住宅の建て方」

		一戸建	集合住宅	その他	計
1人世帯	(実数)	3,461,560	8,340,300	38,600	11,840,460
	(%)	29.2	70.4	0.3	100.0
2人世帯	(実数)	14,937,150	8,419,360	58,800	23,415,310
	(%)	63.8	36.0	0.3	100.0
3人世帯	(実数)	17,331,780	9,036,130	55,480	26,423,390
	(%)	65.6	34.2	0.2	100.0
4人世帯	(実数)	21,580,510	10,099,920	62,280	31,742,710
	(%)	68.0	31.8	0.2	100.0
5人世帯	(実数)	26,358,260	3,779,930	56,240	30,194,430
	(%)	87.3	12.5	0.2	100.0
合計	(実数)	83,669,260	39,675,640	271,400	123,616,300
	(%)	67.7	32.1	0.2	100.0

以上でみてきたとおりである。

ただし、エリア・サンプリングにおける回収率の低さは、エリア・サンプリングという抽出法に基づき実施される調査員調査（面接法や留置法など）によってもたらされるものであり、エリア・サンプリングに固有のものではない。

### 3.2 住宅地図の正確さについて

そこで、何よりも先ず回収率の向上が求められるのであるが、その手立ては調査実施機関ごとのノウハウ・能力に依ることもあり一様ではない。しかし、第一義に求められるのは、抽出された調査対象世帯の住宅が存在しないため調査不能となるような事態を回避するために抽出枠の正確性を確保することである。すなわち、住宅地図の記載情報をチェックし抽出枠の正確性を確保することは必要最小限のこととなる。エリア・サンプリングにおいては、住宅地図を使って調査対象世帯を訪ねていくのであれば、選ばれた世帯が実際にあるかどうかの問題となる。そこで、地図に載っている世帯が実際にあるかどうか、地図上の記載がどの程度正確かを確認する必要があるだろう。

筆者は2008年6月に、東京都調布市の一部の地域において住宅地図の現地照合調査を行った。以下は調査結果の概要である。

対象とした地域は、調布市調布ヶ丘1丁目と4丁目の一部（調布ヶ丘1丁目11～16番地、調布ヶ丘4丁目

13番地、19、20番地、22～31番地、40、41番地）である。地図上の一戸建て住宅（以下、一戸建という）と集合住宅をあわせた住宅数は217戸あった（店舗、事業所等を含む）。調査日は2008年6月20日、使用した住宅地図は2007年11月発行の「ゼンリン住宅地図―調布市」である。実際の照合調査においては、区域を7分割し1地区を2人の調査員により確認した。調査員は電気通信大学の社会調査法演習を受講している学生である。住居環境としては転出転入の変化が比較的少ない住宅地と思われたが、それでも現地においては住宅地図のうちの約1割の変動が認められた。そのとき用いた住宅地図は現地調査の時点では最新のものである。住宅地図の大手であるゼンリンによると、全国の自治体の住宅地図のうち、一部地方では数年おきに作成されるところもあるが、調布市をはじめ都市部では基本的に毎年作成されているという<sup>8</sup>。また、印刷などの発行準備があるため、発行時より少なくとも3か月前には現地の確認が終了していなければならないということであった。したがって今回利用した住宅地図は、短くみても2007年11月より3ヶ月前の実状が確認されているとして、その10ヵ月後にはすでに1割程度の変動が見られたということになる。この間、集合住宅の新築や宅地開発などが行われていたら比べもの

<sup>8</sup> 株式会社ゼンリン広報部に直接問い合わせ聴取による。

にならないほどの変動が生じることにもなる。

地図上の世帯に関しては、一戸建では住宅の形を示した平面図の中に、通常は世帯主と思われる人の名前がフルネームで記載されているが、苗字だけの場合もある。2世帯住宅の場合は、2世帯分の名前が記載されている。もっとも、それは表札や郵便受けにそのように2世帯の表記がされている場合に限られる。住宅地図に名前があるかどうかは、あくまでその世帯が表に名前を出しているかどうかによる。集合住宅の場合も、一戸建の場合と同じく住宅の形を示す図の中に居住者名が記載されているが、スペース上記載できない場合は「別記」に記載されている。その場合はその住宅に別記を示す番号が付され、別記には住宅番号ごとに判明している世帯名が記されている。集合住宅の場合は苗字だけの方が多いが、フルネームで記されている場合もある。

ところで、世帯の中には未成年者だけの世帯があるかもしれない。しかし、表札や郵便受けの表記からはそれを判断することはできないので、そのような稀なケース<sup>9</sup>は無視して、地図上に表示された世帯はすべて調査の対象となる一般世帯と考えることにした。

では、住宅地図に記載されている居住者名が記された世帯数と、現地で居住者名が確認できた世帯数とは、どれだけ差があるのかをみてみることにする。調査員が現地調査を行った7つの地区の記録については、表10および表11のとおりである。

表10はA～Gの各地区ごとに、住宅、世帯、その他の状況について示したものである。まず、一戸建と集合住宅とに分け、一戸建は「居住者名あり」と「居住者名なし」の世帯、「店舗、事業所等」「非居住建物(物置等)」に分類した。集合住宅は「居住者名あり」と居住者名が別記されている「別記あり」、それに「別記なし」の三つに分類し、さらに「別記あり」を「居住者名あり」と「居住者名なし」に分けた。住宅以外としては「駐車場、公園」「その他(建築中など)」がある。各地区ごとに上段は住宅地図における状況、下段は現地調査の状況を示してある。その中で、一戸建と集合住宅ともに居住者名がわかる世帯数が網掛けで示してある。居住者名がわかるこれらの世帯についてみると、まず、一戸建については、地図上で居住者名が記され

ていた世帯は157世帯であったが、現地調査では164世帯の居住者名が確認された。また、集合住宅については、地図上で居住者名が記されていた世帯は151(「居住者名あり」17+「別記あり」134)世帯であったが、現地調査では154世帯(「居住者名あり」17+「別記あり」119+「別記なし」18)の居住者名が確認された。これだけみると、一戸建と集合住宅を足し合わせた居住者名が、地図上では308世帯であったのが現地調査では318世帯となり、3.2ポイントの増加となる。

では、この居住者名について、住宅地図に記載されている居住者名と現地調査により確認できた居住者名の一致状況をみてみよう。表11はA～Gの各地区ごとに、住宅地図に居住者名が記載されている世帯数(住宅地図の世帯数)と現地調査の結果を示している。現地調査の結果については、住宅地図の居住者名と現地で確認された居住者名が同じである「同一」、居住者が替わっていた「変更」、居住者がいるのかわからない、またはいるけれども居住者名が確認できない「不明」、そして、新たに居住者が確認された「新規」に分類した。上段が一戸建で下段が集合住宅である。現地調査の結果、居住者名が確認された世帯の数が「判明世帯数」である。さらに、集合住宅の場合は、「別記なし」の住宅において居住者名が確認された「別記なし判明」を加え、最終的に居住者名が判明した世帯数を「集合住宅小計」とした。そして、一戸建と集合住宅の居住者名が判明した世帯数の計を「合計」とした。

ここでは両者における世帯の一致率を次のようにして求めた。

世帯の一致率=住宅地図と実際とで変わらない世帯の数/(住宅地図と実際とで変わらない世帯の数+変わった世帯の数)= $274/(274+48)=85.1\%$ となる。これは全体の比率であるが、一戸建と集合住宅別にみると、一戸建の「世帯の一致率」= $155/(155+9)=94.5\%$ に対して、集合住宅の「世帯の一致率」= $119/(119+39)=75.3\%$ となる。20ポイント近くの差がある。

ここで、世帯の一致率を求めるときの「変わった世帯」とは、表11でいえば「変更」「不明」「新規」であるが、実際には他にも「空家(空部屋)になっていた」というケースがあり得る。また、一軒の家の中で2世帯に分離していた場合には「新規」となり、逆に2世帯の名前があったが1世帯になっていたという場合には「不明」<sup>10</sup>にした。

<sup>9</sup> たとえば、親戚の子供が同居して、郵便物の受け取りのために郵便受けに別姓の名前が表記されているような場合が考えられるが、そのような事実は外から見ただけでは確認できない。

<sup>10</sup> 2世帯が1世帯になっているケースでは、1世帯が転出したかどうかの確認がとれないので「不明」と

表 10. 調査地区における住宅の状況

地区	一戸建て住宅				集合住宅（マンション、アパート、社宅、事業所ビル等）				駐車場、公園	その他	確認世帯数 注2	居住世帯数の幅				
	居住者名あり		居住者名なし	店舗、事業所等	非居住建物（物置等）	居住者名あり		別記なし								
	住宅数	世帯数				住宅数	世帯数	居住者名なし					居室数	上段住宅数 下段は注1		
			住宅数	世帯数	居室数									世帯数		
A	13	14	0	1	1	1	0	0	3	13	14	9	2	28		
		14											3	4	7	31-44
B	20	23	0	0	1	0	0	0	4	36	36	23	1	59		
	21	25								33	34	26	2	2	2	0
C	21	24	1	1	2	0	0	0	6	8	9	30	2	33		
										6	6	32	0	0	8	30-70
D	27	30	1	0	6	0	0	0	1	6	7	0	2	37		
	28	31											2	2	2	40-42
E	23	23	1	1	2	0	0	0	6	28	38	20	0	61		
	25	25	1	0	1				7	24	30	33		55	1 (建築中)	
F	23	25	3	1	3	0	0	0	2	11	11	1	0	36		
	24	27	2							12	12	0		39	39-39	
G	16	18	0	0	2	7	17	5	17	17	19	19	2	54		
										14	17	22	7	10	3	62-87
地図・計	143	157	6	4	17	7	17	27	119	134	102	102	9	308	1	
現地・計	148	164	5	3	16	7	17	28	108	119	119	119	14	18	22	1

(表中、確認数に変動がなければ空欄にしてある。)

注1) 下段は左から

居住者名ありの居室数  
居住者名ありの世帯数  
居住者名なしの居室数

注2) 上段は住宅地図上の世帯数

下段は確認できた世帯数

表 11. 住宅地図の世帯と現地の世帯の一致状況

上段：一戸建

下段：集合住宅（「合計」を除く）

地区	住宅地図の世帯数	現地調査の結果				判明世帯数	別記なし判明	集合住宅小計	合計
		同一	変更	不明	新規				
A	11	14	0	0	0	14			
	11	12	0	2	1	13	4	17	31
B	23	23	0	0	2	25			
	36	29	4	3	1	34	2	36	61
C	24	24	0	0	0	24			
	9	6	0	3	0	6	0	6	30
D	30	29	1	0	1	31			
	7	7	0	0	0	7	2	9	40
E	23	23	0	0	2	25			
	38	27	1	10	2	30	0	30	55
F	25	24	1	0	2	27			
	11	9	2	0	1	12	0	12	39
G	18	18	0	0	0	18			
	36	29	3	4	2	34	10	44	62
計	157	155	2	0	7	164			
	151	119	10	22	7	136	18	154	318

以上は住宅地図に記載された住宅や世帯についての結果であるが、これらとは別に、住宅地図に記載がなかったが、新たに建てられた住宅（一戸建、集合住宅の別なく）もある。また、今回調査した地域にはそのようなケースはなかったが、地図に住宅の記載はあるが、実際の調査では住宅（建物）がないということもあり得る。

では、これらの「変わった世帯」の内訳はどうか。一戸建・集合住宅別にみると、一戸建の場合は「変更」が2世帯、「新規」が7世帯である。これに対して、集合住宅の場合は「変更」が10世帯、「不明」が22世帯、さらに「新規」が7世帯であった。

A～Gの7地区においては、現地で居住者名が確認できた住宅以外に、表札などがなく居住者不明の住宅

している。

があり、仮にその全住宅に1世帯居住していたとすると<sup>11</sup>、居住世帯の数は、現地で居住者名が確認できた住宅の数（最小）から居住者不明の住宅すべてに居住者がいた場合の住宅の数を加えた数（最大）までの幅があると考えられる。その数は表10右欄外に示したとおりだが、最小、最大それぞれの平均をとると45～66となる。7地区の平均がひとつの国勢調査区の平均世帯数に比較的近いものとなっており、この中のどれかの地区がエリア・サンプリングにおける調査地点としてイメージすることができよう。そこでもう一度、これら7地区の住宅地図で確認した世帯数と現地で確認できた世帯数をみると、(以下の数字は確認できた世帯数で、「住宅地図」→「現地」を示す。) Aは28→31、

<sup>11</sup> 一戸建の「居住者名なし」はすべて空き家であり、住んでいて居住者名不明(表札なし)はなかったため、この数はカウントしていない。

Bは59→61, Cは33→30, Dは37→40, Eは61→55, Fは36→39, Gは54→62となり, その差は-6から+8までの広がりを持っている。

一戸建でも集合住宅でも, エリア・サンプリングにおいてはそもそも住宅地図に居住者名が載っている場合だけ調査対象世帯となり得るので, 新たな世帯の出現は抽出の対象とならない。つまり, 新たな世帯は, 実際には居住者名が分かっても調査の対象から外れるのである。さらに, 集合住宅の場合は「別記」がなく住宅名(〇〇荘, 〇〇マンションなど)しか記されていないものがある。この場合は, 実際にそこに居住者がいても, はじめから調査対象世帯にはならない。また, 住宅(建物)がなくなった, 住んでいる人が引っ越して空家(空部屋)になった, 複数世帯から一部の世帯が除かれたというケースでは, そのような世帯が調査対象世帯になれば調査不能となる。このことは, 表11の「住宅地図の世帯数」が, 現地で調査対象世帯としてどれくらい捕捉されるかをみてもわかる。すなわち, 現地で調査対象世帯となるのは, 現地調査の結果の中の「同一」と「変更」のみである。したがって, 捕捉率は一戸建が100%(157/157)であるのに対して, 集合住宅は85.4%(129/151)である。集合住宅では15ポイントも欠落してしまう。

このように, 住宅地図を用いて行われるエリア・サンプリングでは, 住宅地図の鮮度が調査の精度(回収率)に大きな影響をもたらす。住宅地図の記載状況をもとにそのまま調査対象世帯を特定すれば, 調査員が訪問したとき調査の対象となった世帯の半数が団地造成計画のために転居していたということになりかねない。

以上より, 住宅地図からそのまま調査対象世帯を選んで調査を行ったときには, ①訪問したら空き家になっていたり住宅がない, ②地図上の世帯に新たな世帯が加わっている, などの問題が生じる。調査の前に現状を確認すれば, ③地図上では不明であった住居の居住者名が確認でき, ④地図上にはない新たな住宅の建築による新たな世帯が確認できる, などが可能になる。ただし, 先述の捕捉率の説明からもわかるように, 一般のエリア・サンプリング調査においては, 調査対象となった世帯を訪ね居住者が替わっていても, また表札などで名前が確認できなくても, そこに居住者がいればその居住者を調査の対象としている。つまり, 選ばれた世帯の特定の個人を調査対象者と決めることなく, 選ばれた世帯が住んでいる住居が調査

対象となるのである。

地図情報である以上, 時間的な経過によって実査時には調査地域の一面が一変していることは十分にあり得ることである。地図の改訂版に数年要する地方ではそのような事態はより頻繁に起こり得る。住宅地図が不正確であることによって, その調査地点で本来抽出される調査対象者が抽出されないことになれば, 例えば高齢者の割合が著しく高くなるようなことにもなりかねない。そのことが調査結果の歪みをもたらす一因となる。最終的に調査対象者個人を選定する前に, 世帯の現状に基づいた正確な世帯抽出を行う必要がある所以である。

住民基本台帳から抽出する場合も, 抽出時と調査時で時間差があれば, 僅かながらも転居や死亡などで調査不能となることがある。しかし, 住宅地図の不備により調査不能となることも, 同様にやむを得ないことということはできない。住民基本台帳からの抽出においては, 転居や死亡は事前の確認が不可能である。これに対して, 住宅地図を用いた場合には調査の前に現状との照合が可能である。抽出枠をできるだけ正確な状態で調査を行うことにより, 回収率を高める努力が求められる。

ところで, 現地踏査にはさらに別の問題が存在する。第一に, コストと時間がかかることである。第二に, 現地踏査と実際の世帯抽出作業の関係をどうするかである。

第一の問題については, 調査主体がコストを負担し, 時間の増加についても予め了解することが必要である。調査実施機関がコストを負担したり, 時間的に切り詰めて調査を実施することにでもなれば, 調査の質は間違いなく悪化する。私見によれば, 調査主体は現地踏査のプロセスを仕様で明記し, その実現を経費的にも時間的にも保証しなければならないと考える。

第二の問題については, 抽出における調査員の恣意を極力排除するという基本に従うためには, 調査員が現地踏査を経た住宅地図をもとに, 原則どおり調査機関が世帯抽出を行う必要があると思われる。この手順を示すと, まず最新の住宅地図から調査地点となるエリアを決め(コピーに範囲をマークするなどしたうえで調査員に提供する), 調査員は現地の住宅の実状に合わせて変更があれば地図を訂正し, 調査実施機関へ送付する。それをもとに調査実施機関の担当者が調査対象世帯を抽出して調査員に返送する, ということになる。その際の調査員との連絡方法については, 迅速

な措置が求められる。画像のやり取りが可能なパーソナルモバイルツールを援用するなど、調査実施機関の側でも工夫が求められることになろう。

#### 4. 終わりに

本稿では、世帯における個人の抽出においては、世帯によって個人の抽出確率が異なるためウエイト補正による補正が必要だといわれていることに関して、実験的集計を行い検討した。現在実際に行われているエリア・サンプリングによる世論調査においては、その回収率の低さがもたらす標本の偏りのために、ウエイト補正をすることにより誤差の拡大を生むという懸念があるが、その実相はどのようなものかをウエイト補正の工程を通して確認した。ここでは、現在ほとんど公表されることのないエリア・サンプリングのロウデータを用い、これもまた公表されていない世帯人員別の属性分布という母集団の統計値と対比させながら、ウエイト補正の計算過程を示した。当然のことながら、どの程度の回収率ならウエイト補正をするべきであるとか、どの程度の回収率ならウエイト補正をすべきではないというように一般化することはできない。個々のケースで判断すべきであろう。しかし、どのような場合であれ、世帯における対象資格者数によってウエイト補正を行う以上、その構成比が母集団における構成比と異なれば、ウエイト補正によって別の属性における歪みを生むことにもなる。このようなウエイト補正の実相を踏まえて、それに代わるべきウエイト補正の方法やサンプリングの方法自体が検討されるべきであろう。

次に、抽出枠の正確さと回収率の向上を実現するためには、何よりも住宅地図の鮮度が問題である点を指摘した。そして、地域によっては短期間にかなりの居住者の出入りが行われている可能性があることを確認することができた。そのような現状に対して、調査の前に住宅地図に掲載されている世帯の確認をするべきことと、そのための時間とコストを調査設計に組み入れておくことの必要性を指摘した。

本稿における問題提起をもとに、今後、エリア・サンプリングにおけるウエイト補正をめぐる更なる検討が行われることと、回収率の向上をはじめエリア・サンプリングによる調査の精度を高めるための活発な議論が行われることを期待したい。特に回収率の向上は、対象者とのコンタクトがどのように行われるかによ

て大きな影響を受ける。そのような実査上の問題も重要であるが、何よりも基盤となる抽出枠の整備こそ重要となることを銘記すべきである。

また、近年、住民基本台帳法の改正や個人情報保護法の施行など、調査を実施する者にとっては困難な状況が生まれている。このような状況の下では、調査に携わる関係者間での積極的な情報交換と議論が必要である。特に、調査実施機関はフィールドの現状についての情報交換や協力したノウハウの構築が求められる。本稿で紹介したように、調査機関から実験調査の結果報告がなされたりデータの提供が行われることがなかったら、議論の展開もノウハウの蓄積も行われまい。最後に、これらの調査機関が極めて限られた機会の中で、実験調査の結果を開示したりデータの公表を行っていることを高く評価したい。

#### 謝 辞

本稿に対し、丁寧な査読を通して貴重なご指摘をいただきました審査員のお二人の先生に深く感謝いたします。

エリア・サンプリングの手法と問題点に関しては、本学会の研究部会である社会調査研究会の2008年度研究テーマのひとつであり、研究会における会員諸氏による資料の提供や議論は本稿をまとめるうえでおいに参考になりました。ここに感謝いたします。

ウエイト補正の分析に関して、統計数理研究所名誉教授の鈴木達三先生に貴重なアドバイスとご指導をいただきましたことを、ここに感謝いたします。

本文中、世帯人員別データに用いた2000年国勢調査の結果については、(独)統計センターに依頼し集計をしていただきました。ご協力を感謝いたします。

#### 参 考 文 献

- 朝倉真粧美・小柳雅司・染谷保幸・桜井 薫・武藤大輔 (2004). ランダムテーブルを使用した個人抽出法. 日本行動計量学会第32回大会発表論文抄録集, 310-311.
- 朝倉真粧美・桜井 薫・染谷保幸 (2005). サンプリング方法の違いが調査結果に及ぼす影響. 日本行動計量学会第33回大会発表論文抄録集, 252-253.
- Groves, R.M. et al. (2004). *Survey Methodology*. John Wiley & Sons, Inc.
- JMRA 調査技術研究部会 (2007). 非名簿フレームによる無作為抽出法研究. JMRA (日本マーケティング・リサーチ協会).

- 上村修一 (2008). 住民基本台帳法改正後の動向—会社社アンケートの結果から—。よろん, 101号, 28-33.
- Kish, L. (1949). A procedure for objective respondent selection within the household. *Journal of the American Statistical Association*, vol. 44, 380-387.
- Kish, L. (1965, 第2版1995). *Survey Sampling*. John Wiley & Sons, Inc.
- 松田映二 (2008). 社会調査の課題—世論調査の現場から—。社会と調査, 創刊号, 89-95.
- 根本理恵子 (1983). 住宅地図によるエリアサンプリングの試み。マーケティング・リサーチャー, No. 33, 26-35.
- 小柳雅司 (2003). 住民基本台帳を使用しない個人サンプリング方法の研究。よろん, 93号, 16-19.
- 鄭 躍軍 (2002). 抽出台帳が利用できない場合の確率標本法—意識調査における非標本誤差について—。第70回日本統計学会講演報告集, 382-383.
- 鄭 躍軍 (2007). 抽出枠がない場合の個人抽出の新しい試み—東京都における意識調査を例として—。統計数理, 第55巻, 第2号, 311-326.
- 土屋隆裕・前田忠彦・中村 隆・坂本慶行 (2007). エリアサンプリングによるポスティング調査—郵送調査との比較実験—。統計数理研究所研究リポート 96.
- 氏家 豊 (2007). 世論調査におけるサンプリング。エストレーラ, 6-13.

(2009年6月3日受付, 2010年2月9日最終修正)