



## 国勢調査

宇南山 卓  
(京都大学教授)

### 1 『国勢調査』は社会のインフラ

『国勢調査』とは、総務省統計局が所管する「国内の人及び世帯の実態を把握し、各種行政施策その他の基礎資料を得ることを目的とする」調査で、日本国内に「常住している者は、外国人を含めてすべて」を対象としている。1920年以降、西暦の末尾が0または5の年に実施されており、2020年の調査で100年目を迎えた歴史ある統計である。毎回、予算規模にして約700億円、調査員70万人を動員して実施される、日本で最大の統計調査である。

『国勢調査』は、日本の統計の体系において特に重要な統計である。統計を作成するための調査には、調査対象となる集団（母集団）の一部を調査する「標本調査」と、母集団全体を調査する「全数調査」がある。『国勢調査』は日本に住む個人を母集団とした唯一の「全数調査」であり、標本調査で無作為調査をするための基礎情報（サンプリングフレームとよばれる）を提供する統計的インフラとなっている。

標本調査で母集団の一部だけを調査するのは、より簡易かつ低コストで調査を実施するためであるが、一部の対象だけで母集団全体の動向を正しく把握するには、標本を無作為抽出（ランダムサンプリング）する必要がある。無作為抽出とは、母集団内の各主体が調査対象となる確率が等しいサンプルの選び方であり、適切に標本を抽出するには母集団の全体像の把握が必須である。『国勢調査』がなければ『家計調査』や『労働力調査』など個人を対象とした標本調査も不可能である。

もちろん『国勢調査』は他の統計のためだけでなく、それ自体が学術的にも重要な情報源である。地域別、年齢別、性別などの人口は、社会・経済の分析に不可欠であり、『国勢調査』のみで利用可能である。学術にとどまらず、衆議院の選挙区の区割りの基礎資料となるという重要な役割もある。全国を網羅しつつ詳細な地理的単位での有権者数を把握できるのは『国

勢調査』ならではである。その意味では、『国勢調査』は民主主義にとっても不可欠なインフラである。

### 2 『国勢調査』の調査事項

全数調査では、文字通り母集団を全て調査することが重要になる。しかし、統計調査は強制的に事実の報告を求めるものであり、調査への完全な協力を求めることは難しい。調査の根拠となる統計法という法律では、特に重要な統計（基幹統計）には回答を義務づけており罰則もあるが、現実には強制力を行使することは困難で、全数調査という目的を達成するには、調査内容を限定するなど負担の抑制が不可欠である。

そのため、『国勢調査』の調査頻度は低く、調査事項も他の統計と比べて少なめで、拒否感の強い事項もない簡素な調査となっている。個人ごとに年齢・性別・出生年月・配偶関係・国籍・就業状態、世帯ごとに世帯人員数、住宅の広さ・建て方・保有状況などが調査される。西暦の末尾が0の年は「大規模調査年」とよばれ、調査事項が追加され、学歴などの教育に関する事項や通勤・通学の状況も調査される。

調査結果は、年齢、性別、職業などでクロス集計され、「国勢調査報告」と題した報告書として刊行されてきた。現在では、1985年以降については（1980年は一部）、政府統計のポータルサイトであるe-Stat (<https://www.e-stat.go.jp>)でも同等以上の情報が公表されている。1975年以前についても報告書の一部がPDFファイルで公開されており、ほとんどの情報は電子的に入手可能である。ただし、最新の2020年の結果は現在集計中である。

### 3 『国勢調査』の使い方

#### (1) 属性別の人口比率

『国勢調査』で把握される個人の属性のうち、学術的にも政策的にも最も重要なのは「年齢」である。幼年者は学校に通い、成人後には就労をし、婚姻や出産を経験し、一定の年齢で引退をして老後を過ごすとい

のように、個人の行動はライフステージごとに大きく異なる。各ライフステージにある個人の割合、婚姻や出産を選択した個人の割合などを観察することはさまざまな分析の第1歩である。

たとえば、総人口のうち65歳以上の高齢者の割合である高齢化率、50歳時点での未婚率である生涯未婚率などの計算は、『国勢調査』の典型的な利用例であり新聞等でも報道される。さらに、そうした変数の地理的な分布も非常に詳細なレベルでの集計が可能である。高齢化率や未就学児童数などは介護施設や保育所などのニーズを左右する指標であり、全市町村で信頼できる数値が計算できることは大きな強みである。

さらに、市町村レベルよりも詳細な町丁・字レベルでの集計や緯度・経度に基づき地域を隙間なく網の目（メッシュ）の区域に分けた「地域メッシュ」での情報も公表されている。こうした情報は、商圏人口の計算などマーケティング目的にも使われ、「立地」に関する経済分析の基礎的な情報となる。

#### (2) コーホートデータとしての『国勢調査』

個人のライフステージによる行動の変化を分析するには、同一個人を追跡したパネルデータが必要となる。しかし、日本ではパネル構造を持つデータは、厚生労働省の『21世紀縦断調査』、経済産業研究所の「くらしと健康の調査」(JSTAR) および慶應義塾大学の「日本家計パネルデータ」などに限られている。

この問題に、『国勢調査』を使って対応することができる。『国勢調査』の5年ごとの調査は別体系で調査され個人の名寄せはできないが、原理的には(国際間移動をした者を除き)ほぼ全ての個人を継続的に調査しているため、同一個人の「集団」であれば追跡が可能である。たとえば、『国勢調査』の年齢別集計を整理することで、生まれ年が同一の個人の集団である「生年コーホートデータ」が構築できる。たとえば、宇南山(2011)では、コーホート別年齢別に、女性の未婚率と労働力率を計算し、結婚と就労の両立可能性がどのように変化してきたかを分析している。これは通常はパネルデータでなければできない分析であり、同様の分析はAsai, Kambayashi and Yamaguchi(2015)でも活用されている。

#### (3) 労働力の統計としての『国勢調査』

『国勢調査』の調査事項は限定され、経済学的に分析可能な変数は限られるが、個人の就業状態に関する情報については重要な情報源となる。個人の労働力状態(働いているか)・雇用状態(雇われているか)・従

事する産業・職業などを調査している。

就業状態については、標本調査として『労働力調査』や『就業構造基本調査』などが存在している。また、雇用者(企業に雇われる労働者)については企業側の統計である『賃金構造基本統計調査』『毎月勤労統計』などでも情報が把握できる。

一般に、労働関連の統計は、他の標本調査と比べサンプルサイズが大きい。たとえば、『就業構造基本調査』は100万人程度を対象としている。しかし、『国勢調査』は約1億人の15歳以上人口を全数で調査しており、はるかに詳細な集計が可能である。

たとえば、日本では外国人は依然としてマイノリティであり、その実態を捉える統計は多くない。『国勢調査』であれば比率が小さくとも絶対数は一定程度存在しており、有効な分析ができる。町北(2015)では、外国人労働者の年齢別・学歴別・地域別の動向を分析しているが、これは『国勢調査』ならではの分析と言える。

また、図1は就業者の産業別の割合を示したものである。自営業者やフリーランスなど非雇用者についても情報が把握でき、全数を調査しつつ詳細に分類できるため、正確な変化を捉えることができる。たとえば、新型コロナウイルス感染拡大に伴い、ネット通販や料理の宅配サービスなどが急拡大している。しかし、宅配業者や配達員などは請負契約となり雇用者としての把握はできていない。絶対数は必ずしも多くないため、通常の統計では地域的な偏りや就業者の年齢や学歴別の特徴を把握することは難しい。2020年の『国勢調査』の結果が公表されれば、そうした「ギグワーカー」の実態を明らかにできる。

#### (4) 転居に関する情報

他の統計では利用できない変数が「5年前の住居の所在地」である。5年前に住んでいた場所をたずねた質問から作成され、国内では都道府県・市町村別、国外からの転入については国別に集計される。

たとえば、5年前と現在の住所を比較すると、人口の約2割が5年以内に引越し(住所の変更)を経験していること、引越しをした人の割合は2000年には28%だったのが2015年には22%と減少してきていること、国外からの転入者はそれほど増えていないことなどが分かる。2020年の結果はまだであるが、新型コロナウイルス感染拡大に伴い、外国人の入国制限があった一方で、リモートワークの普及などによる転居などが発生した。こうした現象を正確に分析すれば、コロナ禍での人口移動の実態が把握できる。

(5) 通勤・通学に関する情報

他にも「従業地・通学地別の集計」がある。『国勢調査』では、仕事をしている場所や学校の場所などを「従業地・通学地」と呼んでおり、昼間に過ごす場所とみなしている。住所のある場所ごとの人口を「夜間人口」とよび、そこから従業・通学のために他の地域へ移動してしまう人数を引いて、他の地域から移動してくる人数を加えたものを「昼間人口」とよんでいる。

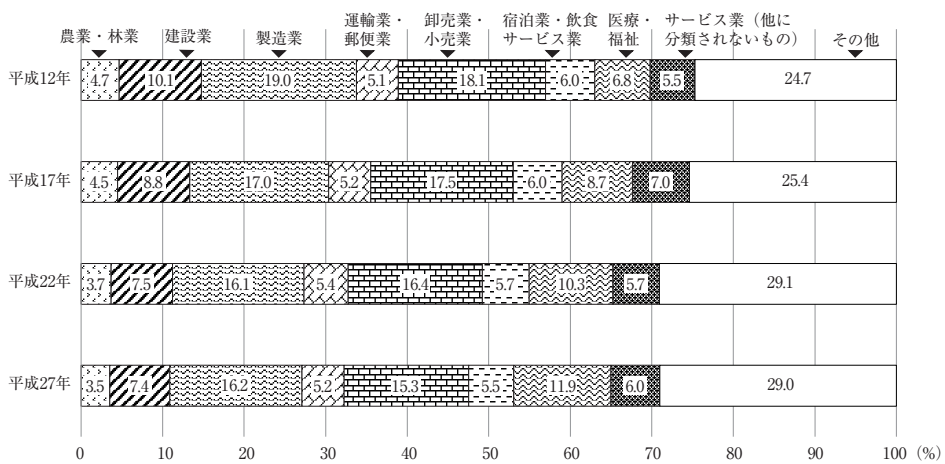
昼夜での他の地域との人口移動を把握することで、繁華街やオフィス街と住宅地などを区別し、社会的なインフラの整備のための基礎資料となる。これは、全数調査だからこそ可能な集計で、統計調査としては他

にない情報である。

図2は、20大都市の昼夜間人口とその比率を示したものである。昼間人口が夜間人口よりも大きいということは、通勤・通学で非居住者が流入してくることを意味しており、いわゆるビジネスエリアであることを示している。政令指定都市という大規模な都市の中でも、昼夜間人口比率が100以上（昼間の方が人口が多い）都市は一部であり、神奈川の3政令指定都市はむしろベッドタウンであることが分かる。

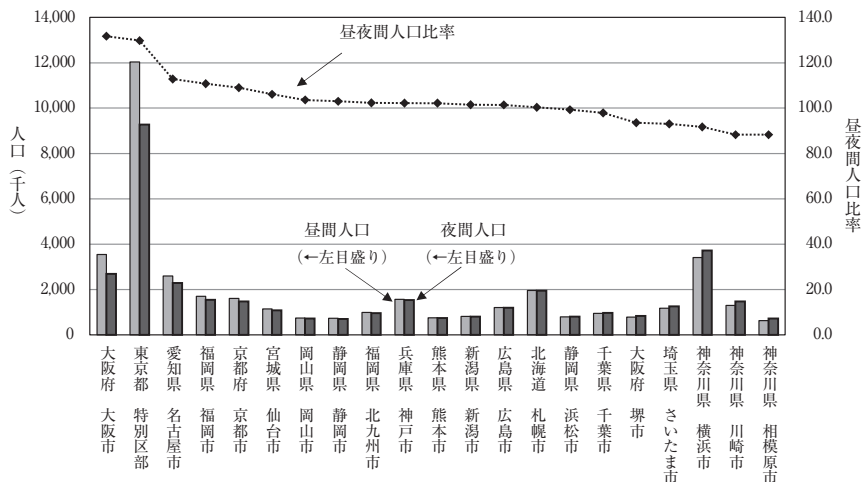
昼夜間人口は、コロナウイルスの感染拡大によるオフィスの分散化などの影響を把握する上で有用な情報となる。個人の位置情報をリアルタイムに把握するようなデータが急速に発展してきている。いわゆる「人

図1 産業（大分類）別 15歳以上就業者の割合の推移



出所：『平成27年国勢調査』就業状態等基本集計結果・結果の概要

図2 20大都市別の昼夜間人口



出所：『平成27年国勢調査』「従業地・通学地による人口・就業状態等集計結果・結果の概要」

流データ」とよばれるもので、携帯電話の通信状況やスマートフォンの位置情報などから、地点別・時点別の人口を把握したデータである。

『国勢調査』は、スマートフォンなどを持っていない個人も対象だが、5年に一度しか調査されず時間別の実態を把握できないため、たとえばコロナウイルスの感染拡大の局面での人口密集状況の分析などはできない。今後も、リアルタイムの人流データの妥当性を確認するためのベンチマークとしての役割が期待されるが、人流データの利用が進めば従業地・通学地の集計の重要性は低下する可能性もある。

#### 4 『国勢調査』における不詳

『国勢調査』を利用する際に注意が必要なのは、近年、調査環境の悪化等により調査の内容が「不詳」になっている者の数が増加していることである。主要な属性についてすらも回答が得られない場合があり、集計表では当該項目が「不詳」と記録される。『国勢調査』の最大の強みは全数調査であるが、「不詳」の存在でこの強みの一部が失われてしまう可能性がある。

『国勢調査』における「不詳」は、調査票が回収されない場合、もしくは回収された調査票に不備がある場合に発生する。総務省の公表する「令和2年国勢調査実施状況（実査編）」によれば、2020年『国勢調査』の調査票の回収率は80.2%である<sup>1)</sup>。調査票が回収できない場合でも、市区町村による照会・聞き取り・補記、住民基本台帳などの行政情報で補完されることもある。こうした作業によって、全人口を把握するという目的は達成できるが、就業状態などのより詳細な情報については把握できない。

項目別の「不詳」の割合を、調査年ごとにまとめたのが表1である<sup>2)</sup>。「不詳」の割合は一般に上昇傾向ではあるが、特に2010年以降に急増している。年齢や国籍が不詳である者は依然として無視できるが、労働力状態不詳の割合が特に高く2015年時点ですでに5%を超えている。また、「不詳」の発生は地域性が強い現象でもある。都市部では、調査票の回収が難しく、近隣からの聞き取りも難しい。表1では東京都の結果も示しているが、どの項目でも全国平均よりも高く、労働力状態は15%を超えている。

「不詳」が存在している場合に、多くのケースで「不詳を除く合計」という処理がされる。たとえば、未婚率を計算するのに、「未婚者」を分子に「婚姻状態不詳を除く総人口」を分母として計算する。こうし

表1 『国勢調査』における不詳の割合

(単位：%)

全国	年齢不詳	婚姻状態不詳	国籍不詳	労働力状態不詳	総人口(人)
2000	0.18	0.78	1.19	1.37	126,925,843
2005	0.38	1.15	0.38	2.63	127,767,994
2010	0.76	1.62	0.82	4.85	128,057,352
2015	1.14	2.13	0.83	5.67	127,094,745
2020	2.32	3.66	1.75	—	126,146,099

東京都	年齢不詳	婚姻状態不詳	国籍不詳	労働力状態不詳	総人口(人)
2000	0.39	2.02	2.08	4.08	12,064,101
2005	1.28	3.40	1.28	8.49	12,576,601
2010	1.44	4.97	1.65	12.09	13,159,388
2015	1.91	6.82	1.39	15.89	13,515,271
2020	3.05	8.96	2.36	—	14,047,594

た計算は「不詳」の人が「不詳」以外の人と同じ属性であることを暗黙に仮定するが、実際にその仮定が成立するとは限らない。

不詳の増加に合わせたデータ分析の必要性が高まっている。『国勢調査』は唯一の全数調査であり、補正することは容易ではない（たとえば宇南山（2013）を参照）。それでも他の統計で代替できない統計であり、注意をしつつ活用をしていく必要がある。

1) <https://www.stat.go.jp/info/kenkyu/kokusei/yusiki32/pdf/08sy0100.pdf>

2) 山田（2002）では2000年調査、小池・山内（2014）では2010年調査における「不詳」を議論している。

#### 参考文献

宇南山卓（2011）「結婚・出産と就業の両立可能性と保育所の整備」『日本経済研究』No. 65, pp. 1-22.  
 ——（2013）「仕事と結婚の両立可能性と保育所——2010年国勢調査による検証」RIETI Discussion Paper Series 13-J-039.  
 小池司朗・山内昌和（2014）「2010年の国勢調査における「不詳」の発生状況——5年前の居住地を中心に」『人口問題研究』第70巻、第3号、pp. 325-338.  
 町北朋洋（2015）「日本の外国人労働力の実態把握——労働供給・需要面からの整理」『日本労働研究雑誌』No. 662, pp. 5-26.  
 山田茂（2002）「2000年国勢調査結果の精度について」『人口学研究』No. 31, pp. 80-84.  
 Asai, Yukiko, Ryo Kambayashi and Shintaro Yamaguchi (2015) "Childcare Availability, Household Structure, and Maternal Employment," *Journal of the Japanese and International Economies*, Vol. 38, pp. 172-192.

うなやま・たかし 京都大学経済研究所教授。著者に *Introduction to Japanese Household Surveys*, Springer (2018年)。日本経済・家計行動・経済統計が専門。