

## Web 付論. 国勢調査における小地域の区分と t-4 年調査対象調査区に対応する t+1 年国勢調査の調査区の設定

### 1. 国勢調査における小地域の区分<sup>1</sup>

国勢調査の個々の世帯は、都道府県・市区町村番号（6桁）、調査区番号（主番号4桁、及び後置番号1桁）、調査区の単位番号（2桁）と世帯番号（3桁）で識別され（これら後置番号を除外した15桁の番号の組み合わせに重複はない）、世帯員番号（3桁）を加えると個人も識別される。一方、もっとも小さな地域の単位は、市区町村別の基本単位区番号（9桁）と調査区番号（主番号4桁）の組み合わせである。基本単位区は、調査区を画定する際の基礎単位とするとともに、調査結果の集計上の恒久的かつ最小の地域単位とすることを目的として1990年国勢調査において導入された。基本単位区は固定的な領域として恒久化されており、4つの要件<sup>2</sup>に該当しない限り将来においても変更せず、また修正にあたっては総務省統計局長への事前の届出が必要になっている。これに対し、調査区は調査員各々の担当区域を明確にし、調査の重複・脱漏を防ぐことを主目的として用いられる。具体的には、調査員1人の担当する規模として適当な人口・世帯規模（概ね50世帯、原則として40～70世帯<sup>3</sup>、かつ1調査区の面積は原則として1km<sup>2</sup>未満）であるように設定されるため、各回の調査ごとの設定替えによる変更がありうる。

国勢調査（10月1日）の実施にあたっては、次の手順で調査区を設定し、その範囲を明示した地図を期日（直近の2020年調査の場合、2月1日）までに作成することになっている。まず、前回調査に用いた基本単位区の点検を行う。その上で、特別調査区<sup>4</sup>と水面調査区<sup>5</sup>を設定し、残る一般調査区について原則として1つの基本単位区または2つ以上の基本単位区を組み合わせる1つの調査区とする。ただし、人口が集中している地域等で、調査区の設定基準に照らして1つの基本単位区内に2つ以上の調査区を設定せざるをえない場合には複数の調査区が設定される。また、1棟で1つまたは2つ以上の調査区となり得る規模の共同住宅がある基本単位区において、共同住宅に調査区を設定した残りの区域のみでは調査区設定基準の世帯数を下回る場合には、特例的に、隣接する基本単位区と組み合わせる調査区を設定することが認められる。このように基本的には、基本単位区（を組み合わせる区域）によって調査区は設定されるが、基本単位区が複数の調査区に分割されることも、基本単位区の一部と別の基本単位区を合わせて調査区が設定されることもある。

なお、基本単位区番号（9桁）は、町又は大字単位の一連番号である町字コード（3桁）、町字の分割番号（1桁）、丁目・小字等の連番（2桁）、基本単位区の連番（2桁）、基本単位

<sup>1</sup> 本節の多くは総務省統計局（2022）による。

<sup>2</sup> ①市町村の境界の変更。②街区方式による住居表示の実施。③町・字界の変更等による市町村内の住居区画の変更。④区画整理、地域再開発、水面埋立て、道路・河川等の新設・改修、災害等による地形・地物の著しい変化。

<sup>3</sup> ただし、人口が集中している市街地等においては、1調査区内の人口が200人を著しく下回らないように設定する。そのため、1調査区の世帯数が70世帯を超える場合もある。

<sup>4</sup> 常住者がいない区域又は著しく少ない区域、特別な施設のある区域、及び、おおむね50人以上の単身者が居住している寄宿舍・寮等のある区域を7種に区分したものである。

<sup>5</sup> 港もしくは水上生活者のいる漁港及び河川の河口とその周辺水域に設定される。

区の分割番号（1桁）から成り，上位6桁で識別される地域を「町丁・字等」と呼ぶ。調査区番号（7桁）は，市町村別に1から始まる連番である主番号（4桁），当該調査区の一般調査区と特別・水面調査区の別を示す後置番号（1桁），2つ以上の基本単位区から成る（場合にのみ付けられる）主番号内の基本単位区の連番である単位番号（2桁）から成る。主番号は各回の調査の度に付けられる連番ということになるが，連番を付ける順序には一定の規則がある<sup>6</sup>。単位番号は1つの基本単位区を1調査区とする場合や1つの基本単位区内に複数調査区が設定される場合には付けられないことになっている。そのため，国勢調査における最小の地域の単位は，原則として市区町村別の基本単位区番号（9桁）と調査区番号のうち主番号（4桁）の組み合わせとなり，基本単位区内に複数の調査区が設定された場合には，国勢調査の小地域集計（基本単位区別）においても基本単位区内の調査区別に統計が作成されている。以下では，このような国勢調査における最小の地域単位となる基本単位区番号（9桁）と調査区番号の主番号（4桁）の組み合わせから成る地域を「基本単位区×調査区」と略記する。また，とくに誤解のない限り主番号（4桁）の示す領域を調査区と呼ぶ。

## 2. t-4年調査対象調査区に対応するt+1年調査区の特定

本文Ⅱ章で述べた通り，t年世帯動態調査はt-4年国勢調査の調査区から無作為に抽出され，調査の対象となったt-4年国勢調査の調査区は市区町村別の調査区番号で特定できる。前節で述べた通り，基本単位区は恒久の領域であり，t-4年国勢調査における調査対象調査区に対応する基本単位区番号はt+1年国勢調査においても原則として同一のものになっている。そのため，調査区は各回の調査において割り振られ一連番号の付け直しが行われる性格のものであっても，t年世帯動態調査の対象調査区をt-4年国勢調査の基本単位区として特定できれば，基本的にはt+1年国勢調査の基本単位区番号を用いてt年世帯動態調査の対象となったt+1年国勢調査の調査区（単一とは限らない）が特定できるはずである。逆に，このように考えると，次の2つの状況では基本単位区番号を用いたt-4年調査対象調査区に対応するt+1年調査区の特定が困難となる。

もっとも代表的な状況は，基本単位区の領域を修正する4つの要件に該当する場合であり，市町村境界や町・字界に変更があった場合である。第2に，t-4年基本単位区が複数の調査区に分割されている等で，基本単位区の一部が調査対象になっていて，t+1年基本単位区内の調査区設定がt-4年から変更される場合である。

前者については，「国勢調査小地域（町丁・字等）境界（都道府県別）」を用いて，t-4年国勢調査における調査対象調査区を含む町丁・字等について，t+1年国勢調査における同一領域の町丁・字等の同定を行った。その上で，同定されたt+1年町丁・字等にt-4年調査対象調査区のある町丁・字等が含まれる可能性が高く，市区町村番号か町丁・字等番号が変わ

<sup>6</sup> 調査区地図は，調査区を明瞭に識別できる縮尺で，市区町村を分割して作成されている。市区町村内での分割の順序は，市区町村役場のある地図を起点とし，東側に隣接する地図，南西方向に隣接する地図の順に，市区町村役場のある地図から始まる渦巻状の順になっている。主番号の順序は，市区町村内を分割した地図の順序にしたがうこととされており，1枚の地図のなかに複数の調査区がある場合には，原則として西北端から東北端，次に南西方向へ移動し東へという右回りの波形の順とすることになっている。

っていると考えられる場合、t+1年の市区町村番号と町丁・字等番号をt-4年に置き換える<sup>7</sup>ことで、町丁・字等内の基本単位区×調査区の設定が同じならば基本単位区番号を用いた特定が可能であるようにした。

後者に関連し、国勢調査における基本単位区と調査区に対応関係別の基本単位区×調査区数をWeb付論表1に示した。原則として、1つもしくは2つ以上の基本単位区を組み合わせることで調査区を設定することとされており、1995年国勢調査で設定された調査区のうち94.2%がこのような類型（Web付論表1の類型Aもしくは類型C）であった。しかし、この割合は一貫して減少しており、2020年は81.3%になっている。逆に、1つの基本単位区内に複数の調査区が設定される類型Bや、基本単位区境界をまたいで調査区が設定される類型Dが増えている。世帯動態調査の対象となった調査区についてみても、Web付論表1の対応関係は全体と大きくは変わらず、大部分の基本単位区×調査区による地域においては基本単位区番号によって調査区全体の領域を識別することができるものの、15～18%が類型Bまたは類型Dであり、基本単位区番号だけでは調査対象となった世帯を特定することができない。

Web付論表1. 国勢調査の基本単位区と調査区に対応関係別 基本単位区×調査区の数

対応関係の類型 <sup>(注1)</sup>	総数						世帯動態調査 対象調査区 <sup>(注2)</sup>			
	1995	2000	2005	2010	2015	2020	2000	2005	2010	2015
	実数									
総数	1,804,045	1,791,093	1,860,839	1,912,778	1,954,153	1,987,159	558	591	581	598
A. 基本単位区と調査区が一对一対応	314,536	280,068	285,077	284,212	285,128	287,086	89	78	89	69
B. 基本単位区が複数の調査区に分割されている	93,595	141,740	166,863	184,784	203,324	222,739	60	56	62	62
C. 調査区が複数の基本単位区に分割されている	1,385,096	1,271,656	1,298,317	1,317,952	1,329,173	1,328,176	377	426	395	423
D. その他（基本単位区/調査区が他方の境界をまたぐ）	10,818	97,629	110,582	125,830	136,528	149,158	32	31	35	44
	%分布									
A. 一对一対応	17.4	15.6	15.3	14.9	14.6	14.4	15.9	13.2	15.3	11.5
B. 1基本単位区=複数調査区	5.2	7.9	9.0	9.7	10.4	11.2	10.8	9.5	10.7	10.4
C. 1調査区=複数基本単位区	76.8	71.0	69.8	68.9	68.0	66.8	67.6	72.1	68.0	70.7
D. その他	0.6	5.5	5.9	6.6	7.0	7.5	5.7	5.2	6.0	7.4

資料：「国勢調査」の調査票情報の二次利用による筆者集計。無人調査区を除く。（注1）対応関係の類型Bは当該基本単位区内に別の基本単位区番号に属する世帯がない場合、類型Cは当該調査区内に別の調査区番号に属する世帯がない場合である。（注2）t年世帯動態調査の対象となったt-4年国勢調査の調査対象地区（たとえば「2015」は2019年（第8回）調査に対応）。調査対象は300調査区であるが、1調査区が複数基本単位区から成る場合があるため、総数は300を超える。

<sup>7</sup> k番目のt-4年町長字等の領域を $S_k$ 、j番目のt+1年町丁字等領域を $T_j$ 、 $S_k$ と $T_j$ の重なった領域を $S_k \cap T_j$ 、領域uの世帯数を $H(u)$ 、領域uの面積を $A(u)$ 、領域uに占める領域vの面積を $A(v \cap u|u)$ で表すと、 $A(T_j \cap S_k) > 0$ であるとき $T_j$ は $S_k$ に同定されたと呼ぶ。領域 $T_j$ に占める $S_k$ と重なった領域の面積の占有割合は $A(S_k \cap T_j|T_j)/A(T_j)$ である。t-4年調査対象調査区を含む領域を $S_0$ とすると、 $A(T_j \cap S_0) > 0$ であるような $T_j$ の市区町村番号と町長字等番号のなかに $S_0$ と同じものがない場合で、すべてのkについて $A(S_0 \cap T_j) \geq A(S_k \cap T_j)$ である $T_j$ について、 $A(T_j \cap S_0|S_0)/A(S_0) > 0.01$ であるか $H(T_j) \times A(S_0 \cap T_j|T_j)/A(T_j) \geq 20$ のときに $T_j$ の町長字等番号を $S_0$ の町長字等番号で置き換えた。最後の条件のうち前者は地図の作成精度の向上といった実際には領域変化を伴わない場合を除外するために課す。後者については、人口分布が一様であると仮定した領域 $T_j$ のうち $S_0$ と重なる部分に居住世帯数が20（1基本単位区あたりの平均世帯数は約30）以上のときに対応する。

1つの基本単位区に複数の調査区が設定されるのは人口・世帯規模が調査区設置基準を超える場合であることを思い起こすと、Web 付論表 1 の基本単位区と調査区に対応関係の変化の背景には都市部への人口集中があるものと考えられる。また、調査区番号の付け方には一定の規則があったため、(基本単位区内に複数の調査区が設定されても) 調査年次間においてその並び順は同様になっていると考えられる。

ここでは次のように、t-4 年国勢調査の調査対象地区の t+1 年国勢調査における特定を図る。まず、Web 付論表 1 の類型 A と類型 C の調査区については、t-4 年基本単位区の全体が調査対象となっていることから、基本単位区番号を用いた特定が可能である。ただし、t+1 年に複数の調査区が設定されていた場合、対応する基本単位区の全体を調査対象とするため、t+1 年調査区番号を t-4 年調査区番号に揃えた。Web 付論表 1 の類型 B と類型 D については、調査区番号を使わずに基本単位区の中なかから調査区を特定しなければならないが、基本単位区内の調査区分割の仕方は変化していない<sup>8</sup>ことを前提に、調査区内に居住する世帯が住んでいる住宅の種類や建て階数、借家住まいが多いといった調査区の特徴を用いる。すなわち、調査区をその特徴に応じ一定の順序に並べ替え、基本単位区内の調査区の並び順を t-4 年と t+1 年で揃えたうえで、基本単位区番号とその中の調査区の順序をキーとして t-4 年調査対象調査区に t+1 年基本単位区×調査区を突合せた。これらの作業で利用した変数は Web 付論表 2 にまとめた。

**Web 付論表 2. t 年世帯動態調査の調査対象となった t-4 年国勢調査区を t+1 年国勢調査において特定するための変数**

変数名	内容
基本単位区番号	市区町村別に設定された9桁の符号であり、上位6桁は町丁字等番号で市区町村と町丁・字境界の変更を補正済み
基本単位区と調査区に対応関係	基本単位区と調査区が一一対応、1つの基本単位区が複数の調査区に分割されている、1つの調査区が複数の基本単位区に分割されている、その他
後置番号	一般調査区、社会施設・大きな病院がある区域、概ね50人以上の単身者が居住している寄宿舎・寮等の区域、水面調査区
基本単位区内の調査区の順序 <sup>(注1)</sup>	基本単位区×調査区別の世帯の半分以上が共同住宅に居住しているか否かの別に、(共同住宅居住者が多い場合) 共同住宅の建て階数(最頻値)、一戸建てを除く住宅の種類(最頻値が民間借家等、公営賃貸、持ち家の順)、持ち家世帯割合(10階級)、共同住宅の居住階数(最頻値)、調査区番号の順にソート、(共同住宅居住者が少ない) 基本単位区×調査区別総世帯数と共同住宅居住世帯数並びに総人口、調査区番号でソート

出典：筆者作成。(注 1)住宅の種類は本文表 2 と共通。

基本単位区内の調査区の順序は、基本単位区×調査区別の世帯の半分以上が共同住宅に居住しているか否かの別に設定した。まず、共同住宅居住世帯が少ない場合、ほとんどが一戸建ての持ち家に居住する世帯の多い、あまり特徴のない住宅地となっている。この場合には、最も基本的な地域(基本単位区×調査区)別の属性である総世帯数と(総世帯数が同じ

<sup>8</sup> たとえば、t-4 年基本単位区が2つの調査区に分割されている場合、t+1 年基本単位区も2つに分割され、かつ t-4 年に同じ調査区であったすべての世帯が t+1 年も同じ調査区に含まれることを前提にする。

場合には) 元の調査区の並び順を用いた。

一方、 $t$  年世帯動態調査の対象となった  $t-4$  年国勢調査の調査区に対応する  $t+1$  年国勢調査の調査区の特定が困難な Web 付論表 1 の類型 B や類型 D は都市部に多いため、共同住宅居住者が半数を超える場合が多い。この場合には、共同住宅の属性(や居住世帯の属性)を調査区を識別する情報として積極的に用いる。まず、建物の階数は客観的に測定可能で誤差が少なく、容易には変化しないため基本単位区内の調査区を構成する世帯の範囲を特定する蓋然性が高い。居住する共同住宅が持ち家である世帯が最も多いのか借家なのかは建物により明瞭に分かれる場合が多いため、住宅の所有関係は世帯属性というより建物(や調査区)の性格を表す属性と言える。そして、中高層共同住宅に複数の調査区が設定される場合、階数で区分される場合も多い。これらの変数の最頻値で調査区を並べ替え、すべてが同一の場合は元の調査区番号の並び順を用いた。なお、基本単位区のなかに複数の特別調査区が設定されることは稀であり、少なくとも世帯動態調査の調査対象となった調査区にはなかったため、後置番号は建物を識別する情報として利用した<sup>9</sup>。

以上のように、Web 付論表 2 の変数を用いて、基本単位区×調査区を単位として  $t-4$  年国勢調査の調査対象調査区に  $t+1$  年国勢調査を突合せた。この方法では、 $t-4$ ~ $t+1$  年の間に基本単位区境域が変化している、あるいは基本単位区内の調査区割が変わっている場合には対応できない。その結果、世帯動態調査の調査対象となった 300 調査区のうち、 $t+1$  年国勢調査の調査区を突合せることができなかった 2009 年(第 6 回)調査の 5 調査区、2014 年(第 7 回)の 5 調査区、2019 年(第 8 回)の 6 調査区を、本文表 3 の手順 2 以後の分析からは除外する。

2010 年国民生活基礎調査について同種の分析を行った 国民生活基礎調査の非標本誤差の縮小に向けた研究会(2018)では、基本単位区×調査区番号をキーとする突合を行い分析対象としたのは約 70%の調査区であった(本文第 II 章)。そこでの結果と比べると、調査区番号を用いない本稿の手法では分析に含めることのできる調査区(約 98%)を大幅に増やしたと言える。一方、この 30%ポイント近い差の背後には、Web 付論表 1 の類型 B や類型 D にあたる  $t-4$  年調査区で、実際には基本単位区内の調査区分割の状況が  $t+1$  年までに変化したにもかかわらず、基本単位区内の  $t+1$  年調査区を  $t-4$  年調査対象調査区と同区域であると誤って特定している場合もあると推察される。これらに対しては、基本単位区内の共同住宅の建て階数や所有状態、居住階数といった変数を手掛かりに  $t-4$ ~ $t+1$  年調査区への対応を個別に精査すれば修正できる可能性があるものの、機械的に処理することが本稿の分析手法の特徴であり、個別処理は行っていない。

## Web 付論 参照文献

国民生活基礎調査の非標本誤差の縮小に向けた研究会(2018)「国民生活基礎調査の非標本

<sup>9</sup> 特別調査区では住宅の種類と所有関係が調査されないため利用できない。国勢調査では特別調査区を 1 つの世帯として調査しているが、世帯動態調査の調査対象となった後置番号 8 の特別調査区(50 人以上の単身者が居住している寄宿舍・寮等)は単独世帯の集まりとみなした。

誤差の縮小に向けた研究会 報告書（平成 30 年 3 月）。（2022 年 11 月 19 日アクセス：  
<https://www.mhlw.go.jp/stf/shingi2/0000204385.html>）  
総務省統計局（2022）『令和 2 年国勢調査 調査区関係資料利用の手引き』総務省統計局。