

Ⅲ 労働力調査の標本設計と結果の推定

第8章 調査世帯の選び方

労働力調査は、全国から調査世帯（住戸）を無作為に抽出する標本調査であり、層化2段抽出法を用いている。本章では、標本調査の考え方のほか、標本調査区の抽出、標本調査区内における住戸の抽出、そして標本交代の方法について解説する。

1 標本調査の考え方

労働力調査が明らかにしようとするのは、雇用者数や完全失業者数など、ある属性を持つ15歳以上の総数である。我が国の15歳以上人口は約1億1千万人（平成21年推計）であるが、労働力調査ではその約1/1100の10万人を調べて全体を推計している。このように一部を調べて全体を推計する調査を標本調査という。属性を明らかにしようとしている対象全体を母集団（この場合、就業状態については15歳以上人口）、抽出されたものを標本と呼ぶ。

労働力調査において、標本による推計を行う場合、大都市に住む者しか含まれていない標本や、収入の多い者しか含まれていない標本のように偏りがあつては、全体の正しい推計は行えない。精度の高い推計をするためには、単に10万人集めれば良いというものではなく、標本が母集団の良い縮図になっている必要がある。

客観的に良い縮図を得る方法として、標本を無作為に抜き出す方法がある。すなわち、くじ引きの原理によって一人ずつ選んでいく方法である。この方法の場合、自然に様々な属性の者が含まれるようになる。しかし、1億1千万人から直接無作為に10万人選び出すというのは、そう簡単ではない。まず、くじ引きのためにはくじを作らなければならないのと同様、1億1千万人のリストを作成する必要がある。しかし、そのためには国勢調査と同様の規模の調査が必要であり、しかもそのリストが調査時点における母集団の姿を反映したものであるためには、毎月メンテナンスをしなければならない。これは労働力調査そのものより大変な作業になる。さらに、仮にリストができて抽出が行えたとしても調査の実施は大変である。10万人が全国にばらばらに住んでいるのであるから、相当多くの調査員が必要となるし、一人の調査員が広い地域に住んでいる数人を探し回るということになる。

このようなことから、実際には、無作為抽出という原則は維持しつつも、まず全国で約100万ある国勢調査の調査区のうちから約2,900調査区を抽出し、次に抽出した調査区の中から約15の住戸を選び出して調査を行うという、いわゆる2段抽出法を用いている。この方法は、幾つかの利点を持つ。まず第1段

目の抽出である調査区の抽出においてリスト作りの困難がない。5年に1度行われる国勢調査において調査区の設定が行われており、このリストから必要な数の調査区を抽出すればよいからである。また、調査区は地面の区画であって、その中に住む人間がどう動こうとも変わることはないから、原則としてそのリストの補正をする必要がない。次に、第2段目の抽出で用いられる住戸リストについても、調査区内の住戸のみをリストアップすればよいから、最新のものを作るのは容易である。さらに、実地調査面からみると、国勢調査の調査区は平均約50世帯を含むように設定されており、一人の調査員が担当するのに適当な数になっている。

このように、2段抽出法を用いることには抽出作業や実地調査上の利点が多いが、精度の良さという点では、全国から直接10万人抽出する場合に比べると一般に劣っている。調査地域の数も2,900程度であるし、また、調査区という「かたまり」を抜き出しているため、例えば社会福祉施設だけからなる調査区のような同じ属性の者の集まりが偏って抽出されてしまうこともあり、標本に様々な属性の者が入るようにして良い縮図を得るという観点からは好ましくない要素もある。

そこで、労働力調査では、精度を上げるために様々な手法を用いている。以下、そのような手法の解説も交えつつ、標本の抽出方法を説明する。

2 標本調査区の抽出

第1段目の抽出は国勢調査の調査区の抽出である。これは、データ化された調査区のリストを用いてコンピュータにより行われる。抽出の際には、単純に抽出していくのではなく、次のような工夫をしている。

(1) 調査区の層化

調査区には、会社の独身寮のあるもの、農家世帯の割合が高いもの、サラリーマン世帯の割合が高いものなど、様々なタイプがある。このことは、例えば産業別の就業者数を高い精度で推計しようとする場合、農家世帯の割合が高い調査区がたまたま多く抽出されるなどということが起こらないような工夫が必要であることを示唆する。そこで労働力調査では、調査区をグループ分けしてから抽出を行う、いわゆる層化抽出法を用いている。このグループを層という。

層化抽出法とは、調査区の持つ特性によって調査区を幾つかの層に分けておいて、各層で独立に抽出を行うという方法である。この方法は、良い縮図を得るためには非常に有効である。

例えば、抽出率1/100（標本を100個に1個の割合で選ぶこと。）で調査区を抽出しようとした場合、層化せずに全国から無作為に抽出する場合に比べ、

地域ごとに層化してから、それぞれの地域において1/100の抽出率で抽出した方が地域間のバランスの良い標本が得られるであろう。また、第1次産業の割合が高い調査区、第2次産業の割合が高い調査区、第3次産業の割合が高い調査区というように、事前に調査区を分類することができるなら、各層でそれぞれ抽出率1/100で抽出を行うことにより産業間のバランスを良くすることができるであろう。この方法を採用すれば、たまたま農業人口の割合が高い調査区が多く選ばれてしまうという危険は少なくなるのである。

このような調査区の層化を行うためには、調査区に関する情報が必要であるが、幸いなことに5年に1度行われる国勢調査の結果から調査区に関する詳細な情報を得ることができる。国勢調査においては、労働力調査に限らず様々な標本調査に使えるよう調査区を分類した資料を作成している。労働力調査では、各調査区の産業、従業上の地位による特性に着目して分類を行ったものを層化基準として作成し、利用している。これは表8-1に示すとおりである。

労働力調査の実際の層化は、この層化基準と、地域区分^{注)}も加味して行っている。地域別に層化を行うというのは、労働力調査の地域別(10地域)表章において、一定の精度を確保することを考慮したものである。

また、本章の4で述べるような標本の交代方法からくる制約から、一つの層から8の倍数の数の抽出を行う必要があり、各層から最低8調査区は抽出するようになっている。このため、地域別に表8-1の分類を行っては、調査区を抽出しようとする上で調査区が不足する可能性がある。そこで、これら分類上の産業特性が類似している層を幾つかまとめて一つの層とし、そこから抽出を行っている。

なお、国勢調査の結果を利用した調査区の分類は、あくまで国勢調査時の情報によるものであるから、国勢調査時から月日がたつと次第に実態から離れたものになってしまう。このため、5年に1度、国勢調査の調査区関連資料がそろった段階で、新しいものに切り替えている。つまり、直近の国勢調査で設定された調査区を新たに層化し、これを抽出のためのリストとして使うのである。なお、「付録6 労働力調査層別調査区数一覧」には、層別の国勢調査調査区数と地域別、層別の標本調査区数を示してある。

注) 標本抽出は11地域別に行うが、公表結果は九州と沖縄を合わせた10地域としている。沖縄県は昭和47年の返還以来、追加抽出されているが、結果については一定の精度を確保できるよう、九州と合わせて集計している。

表8-1 調査区の層化基準

層符号		層化基準
大分類	小分類	
01		後置番号が5（刑務所、拘置所などのある区域）、6（自衛隊区域）、7（駐留軍区域）、9（水面調査区）の調査区
02		人口が0の調査区
03		換算世帯数が15以下の調査区
04		後置番号が4（社会施設、大きな病院のある区域）
		後置番号が8（おおむね50人以上の単身者が居住している寄宿舎・寮などのある区域）
		後置番号が4と8以外で換算世帯数中に占める給与住宅に住む一般世帯数の比が0.5以上の調査区
01	01	学生の寮・寄宿舎（ただし、50人以上の世帯）のある標本単位区
	02	病院・療養所（ただし、50人以上の世帯）のある標本単位区
	03	社会施設（ただし、50人以上の世帯）のある標本単位区
	04	後置番号が4のうち、上記のいずれにも属さない標本単位区
	11	寮などに住む建設業の就業者が50人以上の標本単位区
	12	建設業の世帯の比が0.2以上の標本単位区
	21	寮などに住む製造業の就業者が50人以上の標本単位区
	22	製造業の世帯の比が0.3以上の標本単位区
	31	寮などに住む卸売・小売業、飲食店、宿泊業の就業者が50人以上の標本単位区
	32	卸売・小売業、飲食店、宿泊業の世帯の比が0.3以上の標本単位区
	41	寮などに住む金融・保険業、不動産業の就業者が50人以上の標本単位区
	42	金融・保険業、不動産業の世帯の比が0.2以上の標本単位区
	51	寮などに住む電気・ガス・熱供給・水道業、情報通信業、運輸業の就業者が50人以上の標本単位区
	52	電気・ガス・熱供給・水道業、情報通信業、運輸業の世帯の比が0.3以上の標本単位区
	61	寮などに住む医療、福祉、教育、学習支援業、複合サービス事業、サービス業の就業者が50人以上の標本単位区
	62	医療、福祉、教育、学習支援業、複合サービス事業、サービス業の世帯の比が0.4以上の標本単位区
	71	寮などに住む公務の就業者が50人以上の標本単位区
	72	公務の世帯の比が0.4以上の標本単位区
	91	後置番号が8の調査区のうち、上記のいずれにも属さない標本単位区
	92	給与住宅に住む一般世帯数の比が0.5以上の調査区のうち、上記のいずれにも属さない標本単位区
05		漁業の就業者の比が0.2以上の調査区
06		漁業の就業者の比が0.1以上0.2未満の調査区
07		建設業、製造業の業主の比が0.1以上の調査区
08		卸売・小売業、飲食店、宿泊業の業主の比が0.1以上の調査区
09		情報通信業、運輸業、金融・保険業、不動産業、医療、福祉、教育、学習支援業、複合サービス事業、サービス業の業主の比が0.1以上の調査区
10		農林業の就業者の比が0.3以上の調査区
11		農林業の就業者の比が0.1以上0.3未満の調査区
12		公務の就業者の比が0.1以上の調査区
13		金融・保険業、不動産業の雇用者の比が0.1以上の調査区
14		製造業の雇用者の比が0.3以上の調査区
15		建設業の雇用者の比が0.1以上の調査区
16		医療、福祉の雇用者の比が0.1以上の調査区
17		卸売・小売業、飲食店、宿泊業の雇用者の比が0.2以上の調査区
18		教育、学習支援業、複合サービス事業、サービス業の雇用者の比が0.2以上の調査区
19		電気・ガス・熱供給・水道業、情報通信業、運輸業の雇用者の比が0.1以上の調査区
20		製造業の雇用者の比が0.2以上0.3未満の調査区
21		製造業の雇用者の比が0.1以上0.2未満の調査区
22		卸売・小売業、飲食店、宿泊業の雇用者の比が0.1以上0.2未満の調査区
23		教育、学習支援業、複合サービス事業、サービス業の雇用者の比が0.1以上0.2未満の調査区
99		上記のいずれにも属さない調査区

平成17年国勢調査結果による層化基準（平成20年5月に調査開始の調査区から適用）

(2) 確率比例抽出

調査区を抽出する場合、調査区のリストから抽出を行うのであるが、一つずつ無作為に抽出していくのではなく系統抽出法を用いている。系統抽出法とは、調査区に一連番号を付け、番号を等間隔に選んでいく方法である。例えば、200調査区から10調査区抽出する場合（抽出率1/20）、1から200までの番号を付け、1から20までの数字から無作為に一つの数字を選び、仮にそれが7とすると（これを抽出起番号という。）、7に抽出率の逆数（これを抽出間隔という。）を次々に足して得られる数、すなわち7, 27, 47, …… , 187の番号を付けられた調査区を抽出することになる。系統抽出は、作業が非常に容易であるという利点に加え、例えば調査区を市町村ごとに並べておけば一部の市町村に偏ることがなくなり、層化に似た効果も期待できる。

さらに、労働力調査では、系統抽出を行う際に調査区の規模も考慮して精度の向上を図っている。それは確率比例抽出と呼ばれるもので、層化が調査区の産業特性に関する情報を主として活用したのに対し、調査区内の世帯数に関する情報を活用しようとするものである。

国勢調査の調査区は、50世帯を標準に設定されているが、実際には世帯数はかなりばらついている。これは精度を低下させる原因となるが、もし事前に調査区の規模に関する情報が利用できれば精度がかなり改善されるのである。

例えば、3調査区から一つ抽出し、その調査区に居住する者を全て調べて3調査区全体の就業者数を推計する場合を考えてみる。調査区（A, B, C）の規模が次のようになっていたとして、A, B, Cのいずれかを調べて120人という値を推計してみる。

	A	B	C	計
総人口	100人	60人	40人	200人
就業者	60人	40人	20人	120人

無作為抽出や、上のような系統抽出の場合、A, B, Cのそれぞれが選ばれる確率（抽出確率）は1/3である。仮にAが選ばれたとすると、全体の就業者数の推定値は、

$$60 \text{ 人} \times \frac{3}{1} = 180 \text{ 人}$$

ということになる。ここで3/1倍したというのは、抽出率1/3の逆数を乗じたのであるが、抽出確率の逆数を乗じたともいうことができる。同様にしてB及びCが抽出された場合の推定値は、それぞれ120人、60人となる。

さて、仮に調査区の総人口が事前に分かっていたとする。つまり、A、B、Cの人口規模の比が5：3：2であることが事前に分かっていたとする。このとき、A、B、Cから一つ抽出する際にそれぞれ0.5、0.3、0.2の確率で抽出されるようにする。具体的にはAが5枚、Bが3枚、Cが2枚の計10枚のカードから1枚引いて、そのカードがAであったらAを抽出するという方法を採用。Aは10枚中5枚あるからAが選ばれる確率は0.5であり、Bは10枚中3枚だから確率0.3、Cは10枚中2枚だから確率0.2となる。このような抽出法を確率比例抽出という。

このケースでは、全体の推定値は、等確率で選んだ場合と同様に抽出確率の逆数を乗じることによって得られる。例えばAが選ばれた場合は、

$$60 \times \frac{1}{0.5} = 120$$

となり、B、Cが選ばれた場合はそれぞれ

$$40 \times \frac{1}{0.3} = 133.3\dots \quad 20 \times \frac{1}{0.2} = 100$$

となる。確率比例抽出を行っていることにより、例えば10回に5回は120という推定値になる。ここで先の例と比べてみると、A、B、Cから等確率(1/3)で選んだ場合は推定結果が180、120、60と分布しており、これらの推定値はいずれも3回に1回の割合で得られることになるが、確率比例抽出を行った場合はかなりばらつきが小さくなり、しかも母集団の値に近い推定値が得られることが多くなっている。

このように確率比例抽出を行うことにより一般に精度が向上する。この例では調査区の規模が完全に分かっていると仮定したが、ある程度近似的な状況でも同様な効果が期待できる。

このような考え方により、労働力調査では、国勢調査時に得られる情報から換算世帯数を求めて利用している。換算世帯数とは、

$$\text{換算世帯数} = \left(\begin{array}{l} 2人以上の \\ \text{一般世帯数} \end{array} \right) + \frac{(1人の一般世帯数) + (\text{施設等の世帯人員})}{3}$$

注) 施設等とは寮、寄宿舎、病院などをいう。

という算式で表され、世帯員3人の世帯を基準として、何世帯相当あるかを示したものである。実際にはこの情報は次のように集約してウェイトという形で使われている。

換算世帯数	ウエイト
1～15	1
16～30	2
31～45	3
46～60	4
⋮	⋮

実際の抽出に当たっては、ウエイト2のものは2個同じものを並べるというようにして調査区を1列に並べ、そこから等間隔に抽出を行うという方法を採用している。この用法は確率比例系統抽出とも呼ばれている。

3 標本調査区内における住戸の抽出

第2段目の抽出、すなわち抽出された調査区において実際に調査を行う対象を選び出すという作業は、調査区の抽出のように機械の中で行うのではなく、調査員が実地に調査区を巡回してリストを作成し、指導員が抽出を行っている。この場合問題になるのは、何のリストを作って抽出を行うかである。

労働力調査は個人の属性を調べる調査であるから、抽出の方法として、①調査区内に居住する者のリストを作成し、個人を直接抽出する方法がまず考えられる。また、②世帯のリストを作成し、抽出した世帯の世帯員について調査する方法、③人間ではなく建物やアパートの部屋といった「入れもの」のリストを作り、抽出した「入れもの」に居住する世帯を調査する方法なども考えられる。

労働力調査では、このうち③の方法、すなわち「入れもの」のリストを作成して抽出を行うという方法を採用している。この「入れもの」を「住戸」と呼んでおり、抽出の際の単位となるものとして「抽出単位」（調査区を第1次抽出単位とみることの対比でいえば第2次抽出単位）とも呼んでいる。抽出単位（住戸）は「調査区内にある住宅やその他の建物の各戸で、一つの世帯が居住できるようにになっている建物又は建物の一区画」と定義され、例えば、一戸建住宅の場合はその建物全体が抽出単位（住戸）となり、アパート、マンションなど場合は一戸一戸が抽出単位（住戸）となる。

住戸のリストで抽出を行う大きな理由は、リストの老化を防ぐことができる点にある。労働力調査では指定された調査区は4か月継続して調査を行い、リストの作成は開始月の前月に行うことになっている。このとき、仮に世帯や個人のリストを用いるとすると、リスト作成時から調査時点までに転出・転入や死亡などのため、リストの内容が変わってしまい、精度の高い推計ができない。ところが、住戸のリストの場合、人間がどう動こうとも、調査時点で抽出された住戸に住む者を調べればよい。

なお、空き地に急に家が建ったり、急に家が取り壊されたりした場合は、リストの補正を行っている。

以下に示す実際の抽出単位の把握及び抽出作業では、調査区地図と抽出単位のリストがこのリストに当たる。このうち、調査区の境界及び道路、河川、鉄道や建造物など目印となるものを記入するのが調査区地図で、調査区内の全ての抽出単位を把握することができる。各抽出単位の名称や住所、居住者の有無を記入するのが抽出単位のリストで、このリストから抽出単位の選定を行っている。

(1) 住戸の把握

住戸の把握に当たっては、上に述べたような理由から、調査時に人が居住している可能性のあるものは全て抽出単位のリストに入れておかなければならない。居住部分のない事務所や工場は、人が住む可能性がないので把握する必要がないが、空き家は人が住む可能性があるので把握する。また、たとえ建築中の家があっても、調査時に完成している可能性があれば把握しておく必要がある。抽出単位のリストの作成は、正確な調査を行うためには極めて重要な作業である。

なお、病院、高齢者介護施設のような社会福祉施設、建設従業者宿舎などでは、部屋ごとに抽出単位（住戸）としているが、1室が10人以上収容できるようになっている場合、柱や通路などの目印によって更に小さく分割することとしている。これは、精度を考えた場合、各抽出単位内に居住する者の数が均等に近い方が好ましいというのが主な理由である。

(2) 調査を行う住戸の抽出

調査を行う住戸の抽出は、住戸に一連番号を付して、等確率で系統抽出を行っている。一連番号は、把握時に居住者のなかったものから番号を付け、次に居住者があったものに番号を付ける。これは、層化と同じ効果を狙ったもので、この方法により、調査区内における居住者がいない住戸、居住者がある住戸の比に応じて、住戸が抽出されるようになる。

抽出率は、調査区のウエイトの逆数を用いている。ウエイトは、15世帯を単位に定められていたから、例えば国勢調査時に換算世帯数が50であった調査区は、ウエイトは4で、その調査区における住戸の抽出率は1/4となる。このとき、抽出単位のリストで抽出の起点（抽出起番号）を2とすると、 $2 \cdot 6 \cdot 10 \cdots 46 \cdot 50$ の計13世帯が抽出される。この方式を採ると、調査区における抽出単位数が多くなるに従って抽出率は小さくなり、どの調査区も15世帯程度

調査されるようになる。これは、後で述べるように推定式（第9章）が簡単になるという利点と、調査員の事務量が平均化するという利点を持っている。

4 標本の交代方法

雇用・失業動向などをみるために労働力調査の結果を利用する場合、前月差（比）や前年同月差（比）によって動向を把握することが多い。このため、各月の推定値の精度向上のみを考えるだけでは十分でなく、前月、前年同月との比較上の安定性のための工夫も必要となる。

労働力調査では、前月と継続して調査する標本と、前年同月に調査し、再び調査する標本とを適度に織り交ぜることによって前月、前年同月との比較の安定性の向上を図っている。例えば、前月と今月を比較する場合、同じ標本による推定値で前月と今月を比較した結果は、異なる標本でそれぞれ独立に得られた推定値を比較した結果に比べ、より安定した値になっていることが予想される。一方、同じ世帯を長い間調査したりすると、標本に偏りが生じたり、世帯への負担の問題が生じたりするため、標本を完全に固定するのも問題がある。そこで労働力調査では、毎月標本の一部を交代させる方法を用いている。具体的な交代方法は次のとおりである。

- ① 標本調査区は4か月継続して調査し、毎月1/4ずつ新しい調査区に交代する。また、標本調査区は、1年後の同じ時期にも調査を行う。したがって、ある月をみた場合、半分が1年目の調査区、残りの半分が2年目の調査区ということになる。また、一つの調査区に関しては、1年目の4か月、2年目の4か月、合計8か月の調査を行うことになる。
- ② 同じ調査区を4か月継続して調査するので、2か月ずつ前期と後期に分け、前期と後期で違う住戸を調査する。したがって、調査対象となった世帯は、同じ住戸に居住していれば2か月継続して調査され、1年後の同じ時期に再び2か月継続して調査されることになる。

この標本の交代方法を図示すると図8-1のようになる。図で分かるように、ある月をみた場合、調査を行う調査区は8組に分けられる。まず、調査区が今年初めて標本調査区となった調査区(これを「1年目調査区」という。)か、前年同月に調査し、再び調査を行う調査区(これを「2年目調査区」という。)かにより二つに分かれる。次に、調査区を4か月継続して調査するうちの前期か後期か、住戸を2か月継続するうちの1か月目か2か月目かにより四つに分かれ、計8組となる。この8個のグループは、それぞれ独立に全国世帯の縮図になっていると考えられるから、独立に全体の推定値を計算することができる。

この8組の推定値は、第9章で述べる誤差計算に利用されている。

図8-1を見て分かるように、前年との継続を見ると、2年目の調査区が常に半分含まれており、調査を行う住戸のうち半分は前年の同じ月に調査を行っている。また、前月との継続をみると、3/4が同じ調査区となっており、調査を行う住戸のうち半分は前月にも調査を行った住戸である。これらは、それぞれの前年同月との比較、前月との比較の安定性を向上させる効果を持っている。

図8-1 標本の交代方法

組別符号	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
A-1(1年目)	前 1か月目	期 2か月目	後 1か月目	期 2か月目								
A-2(2年目)												
B-1(1年目)												
B-2(2年目)												
C-1(1年目)												
C-2(2年目)												
D-1(1年目)												
D-2(2年目)												

図8-2 標本交代の例

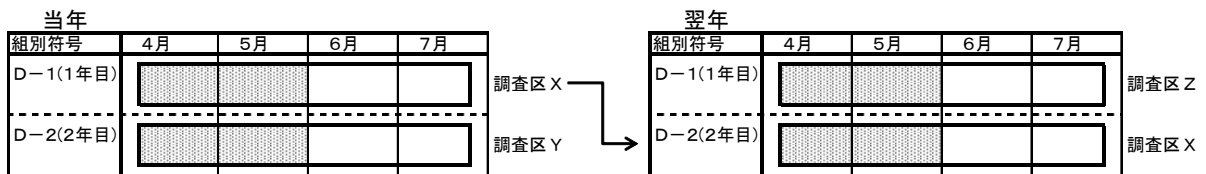


図8-3 例) 5月の調査区

	組別符号			
1年目	A-1	B-1	C-1	D-1
2年目	A-2	B-2	C-2	D-2

	組別符号			
前期	A-1	A-2	D-1	D-2
後期	B-1	B-2	C-1	C-2

	組別符号			
1か月目	A-1	A-2	C-1	C-2
2か月目	B-1	B-2	D-1	D-2

このとき、例えば2年目・前期・2か月目はD-2といえる。