

【BOX】標本設計とサンプル誤差

実際の世論調査では、国内すべての世帯を対象とすべきであるが、費用や時間など様々な事情から、すべての対象に調査を行うことは困難である。そこで、通常は、何らかの方法によって調査対象を抽出し、その結果をもって全体を推測する標本調査を行うことになる。

標本調査の主なポイントは、調査結果ができるだけ「真の世論（国内すべての世帯に調査したときの結果）」に近くなるよう、偏りのない調査対象を抽出すること（標本設計）と、「真の世論」との乖離があるとすれば、それがどのくらいあるのか（調査結果をどのくらいの幅を持ってみるべきか＝調査結果の誤差）を知っておくことである。

標本設計

本調査では、標本設計に「層化2段無作為抽出法」という方法を用いている。この方法では、全国の地域性、都市規模の特性に偏りがなく、無作為（ランダム）に調査対象を選ぶことができる特徴がある。

層化2段無作為抽出法の手順

地域別に調査地点数を按分	全国を9地域（北海道、東北、関東、北陸、中部、近畿、中国、四国、九州）に区分し、各地域の普通世帯数に比例して、全国の調査地点数400地点を各地域に按分する。	（例） 北海道(27地点)
都市規模別に調査地点数を按分	1地域の中で都市規模別6グループ（大都市<全国で13>、世帯数4万以上の市、世帯数2万以上4万未満の市、世帯数1万以上2万未満の市、世帯数1万未満の市、郡部<町村>）に分類し、調査地点数を各グループの普通世帯数に対応させて、各グループに割り振る。	大都市(8地点) 4万以上の市(8地点) 2万以上の市(2地点) 1万以上の市(2地点) 1万未満の市(1地点) 郡部(6地点)
調査地点の選定	グループごとに割り振られた地点数を各グループの中から無作為に抽出し、調査対象地点を決める。	8地点を大都市のグループに属する地域からランダムに選ぶ
調査対象世帯の選定	調査地点から、住民基本台帳に基づき無作為に各15世帯の調査対象世帯を選んでアンケートを実施する。	さらに、その8地点からランダムに各15世帯を選ぶ

調査結果の誤差

本調査の調査対象世帯数は6,000世帯である。例年、回収率は70%前後なので、4,200程度の世帯の回答が得られる。調査の精度は、サイコロを多く転がせば“1”の目が出る確率が限りなく6分の1に近づくと同じように、「標本数（調査対象世帯数）が多いほど、調査結果は真の姿（真の世論）に近づいていく」という“大数の法則”に基づいている。

では、4,200世帯の回答結果は、どの程度の誤差を持っているのだろうか。例えば4,200世帯のうち、Aという考えを持つ世帯の割合が、昨年は60.0%で、本年は60.5%であったとすると、「昨年に比べて本年はAと考える世帯が増えた」と評価してよいだろうか。

下表は、調査世帯数と調査結果の比率に応じた誤差の範囲を示した早見表である。これによれば、先程の結果は、プラス・マイナス2%強の誤差を持っているので、必ずしも「昨年に比べて本年は増えた」とは言えないことになる。

調査結果の誤差（信頼度95%）

調査結果の比率 世帯数	1% (99%)	5% (95%)	10% (90%)	20% (80%)	30% (70%)	40% (60%)	50%
500世帯	1.3	2.8	3.8	5.1	5.8	6.2	6.3
1,000世帯	0.9	1.9	2.7	3.6	4.1	4.5	4.5
2,000世帯	0.6	1.4	1.9	2.5	2.9	3.1	3.2
3,000世帯	0.5	1.1	1.5	2.1	2.4	2.5	2.6
4,000世帯	0.4	1.0	1.3	1.8	2.0	2.2	2.2