

研究員 の眼

標本調査はサンプル抽出が命 何人からアンケートをとればいいか？

保険研究部 主任研究員 篠原 拓也
(03)3512-1823 tshino@nli-research.co.jp

人々が何を考えているのか、ある問題に対しどのような意見を持っているのか、を調べることは重要だ。そのために、公共機関やマスコミや企業は、社会調査や世論調査や市場調査を実施する。このうち、日本で最も規模が大きいものは国勢調査であろう。これは、総務省が5年ごとに行う「全数調査」であり、全ての人及び世帯が対象となる。しかし、全数調査は時間と手間がかかるため、頻繁に行うことは難しい。そこで、日頃、よく行われるのは、ランダムに抽出した一定の人々をサンプルとする「標本調査」である。

標本調査では、時間と手間が節約できるが、全数調査に比べて調査結果の信頼度は小さくなってしまふ。そこで、一定の信頼度を確保するためにはどの程度のサンプルの大きさが必要なのか、即ち、何人からアンケートをとればいいか、ということを統計理論から明らかにする必要がある。なお、話を簡単にするために、以下では、調査の母集団はサンプルの大きさよりも圧倒的に大きいものと仮定する。

まず最初に、注意しておくべきことがある。統計理論では、推定結果を言葉で正確に表そうとすると、表現が、まどろっこしくなるということである。

例えば、現政権の支持率を推定するために、有権者に対して標本調査によるアンケートを実施したとする。1,000人の有権者のうち、500人が「支持する」と回答した。このとき「支持率は50%」と報道される。これは間違いではないが、統計理論の表現としては、やや厳密さを欠く。

厳密には、「支持率は50%と推定され、95%の確率で、誤差は3.1%以下」、といった表現になる。これは、このような標本調査を100回行ったら、そのうち95回は3.1%以下の誤差、つまり、支持率が46.9%から53.1%の範囲内におさまる、ということを主張している。標本調査の信頼度を、確率と誤差で表現しているのである。

それでは、誤差を小さくするにはどうしたらいいだろうか。統計理論上、この例の「95%の確率で、誤差は3.1%以下」という表現は、「80%の確率で、誤差は2.1%以下」と言い換えることができるが、これでは、誤差が小さくなる一方、その誤差におさまらない確率が高くなってしまうため、あまり意味がない。確率を95%にしたままで、誤差だけを小さくするためには、サンプルを大きくする必要がある。

例えば誤差を1%まで小さくするためには、何人からアンケートをとればいいだろうか。統計理論上は、10,000人からアンケートをとる必要があるという結果になる。つまり、アンケートの人数を10倍に増やさなくてはならない。

調査の結果が五分五分ではなく偏ったものになるだろうとあらかじめ予想できる場合は、アンケートの人数を減らすことができる。政権支持率の調査の例で、事前調査の結果等から、あらかじめ支持率は20%程度と考えられる場合、「95%の確率で、誤差は1%以下」とするための人数は6,400人となる。人数を10,000人の、3分の2未満に減らせるわけである。

このような標本調査では、サンプルを大きくして誤差を小さくするとともに、サンプルをランダムに抽出することが大切である。例えば、東京都在住の30歳代の男性会社員にばかりアンケートをして、サンプルを大きくしたとしても、ランダムに抽出したとは言えない。

マスコミの世論調査では、「層化二段無作為抽出法」がよく利用される。この方法では、例えば、全国をいくつかのブロックに分け、それぞれのブロックで、市区町村を、都市規模や産業別就業人口構成比などによって一定数の層に分類する。まず、各ブロックの人口数の大きさに比例して、各層から無作為に調査地点を抽出する。次に、住民基本台帳をもとに、各調査地点ごとに、一定数のサンプルを無作為に抽出する。このようにすることで、ランダムなサンプルの抽出が可能となる。

アメリカの事例だが、昔、あるマスコミがランダムに電話をかけてアンケートを行う形で、世論調査を行った。ところが、この方法では、電話を持っていない低所得者層と、着信を選別してなかなか電話に出ない高所得者層がアンケートから抜け落ちてしまい、中間所得者層ばかりのサンプルになってしまったということである。固定電話が減少し、スマートフォンをはじめとした携帯電話や、ブロードバンド回線を利用したIP電話が拡大を続けている現在の日本では、こうした電話アンケートについても、その方法の影響を考慮する必要があるように思われる。

テレビ局や新聞などのマスコミは、時々、世論調査を行って結果を報道する。企業も、商品やサービスについての顧客アンケートの結果を公表する。このような場合、ついつい、結果にばかり関心が向いてしまう。「このアンケートのサンプルの大きさはどれくらいなのか」「サンプルはどのように抽出したのか」といった調査の方法にも、注意する必要があると思うが、いかがだろうか。