

●What's EBM ? ⑩

「代表的な意見」とは何か？－「237万人」と「2万人」

中山 健夫

I. 大統領になるのは誰か？

今回は情報一般を読む時の基本的な、そして最も大切な考え方の一つをお話します。

1936年、アメリカ大統領選挙で一つの大きな事件が発生しました。選挙前に237万人という膨大な有権者から、「誰に投票するか」の回答を得たリテラシー・ダイジェスト社は、確信をもって共和党候補のカンザス州知事アルフレッド・ランドンの当選を予測しました。大恐慌からの脱出を掲げて、1933年にルーズベルトが「ニューディール政策」を唱えて大統領に就任、その2期目の選挙です。時あたかも、欧州でナチス政権が誕生し、日本は国際連盟を脱退して中国大陸への進出…と国際危機が明らかになりつつある頃です。

巨額の費用を投入して大調査を実施したリテラシー・ダイジェスト社に対し、わずか2万人の調査結果からルーズベルト再選を予想したのがギャラップ社です。「237万人」対「2万人」。さて軍配はどちらに上がったのでしょうか…？

結果はルーズベルトが選挙人の獲得数で523対8と、ランドンに圧勝しました。ランドン有利と予測したリテラシー・ダイジェスト社は社会的信用を失なって結局倒産の憂き目に遭います。一方のギャラップ社は、世論調査の世界で今日に至る名声を勝ち得ていくことは、関係者のよく知るところとなっています。

II. どうして237万人が負けたのか？

リテラシー・ダイジェスト社の失敗の原因は果たして何だったのでしょうか？

それは第一に国民の有権者を探すのに、自社雑誌の購読者名簿、電話帳、自動車登録名簿から1,000万人を選んだことでした。こうして選ばれた人々は全有権者の中で経済的に恵まれた人に偏ってしまいました。実際に調査に回答したのは、そのうちのわずか23%であったことが第2の失敗です。リテラシー・ダイジェスト社は、対象者を選ぶサンプリングで「全国民の代表的意見」を推測するのに不適当な相手を選び、さらにその中からの少数の偏った回答—たとえそれが237万という膨大な数だったとしても、分母である1,000万人から見ると一部に過ぎません—を無神経に一般論にするという二重の過ちをおかしたわけです。このような「不適切なサンプリング」は、情報を読む際の疫学・EBM的な着眼点である「選択バイアス（かたより）」による誤りの代表です。

一方のギャラップ社は、居住地、年齢、性別、人など有権者全体の比率に似るよう工夫して選び出しました。その結果、この「2万人」の情報も、偏った「237万人」と比べられないほど「全国民の代表的意見」を反映したのです。

アンケート調査は手軽に誰でもできるように思われており（それは一面事実ですが）、それが科学的に意味を持つにはどういう配慮が必要か、専門家以外には十分知られていないようです。多人数を調べて、適当な統計解析をすると、良い結果が得られるように錯覚されているような傾向もあります。ちなみに研究者は、アンケー

ト調査という言葉はその印象の「軽さ」のためか、用いることはあまり好まず、「質問票調査（英語では questionnaire survey）」と呼ぶことが一般的です。「数が多ければ良い」という誤解は、アンケート調査に限らず、医療者によるものも含め「研究」と称されるものかなりの部分で見られます。そのような時には、私たちはリテラシー・ダイジェスト社が身を以って示した教訓を思い出す必要があるでしょう。

Ⅲ. 「国民の代表的意見」…?

日本では「国民の代表的意見」というと、話上手な政治家や主義主張の強い特別なグループ、またはマスメディアの意見のように思われがちです。実際は声の大きい人の方が少数派で、多数派の方はおとなしい（サイレント・マジョリティー：物言わぬ大衆と言われます）こともよくあります。このような声なき声をどれだけ把握できるかということが、全体の傾向を知る上で欠かせません。

声を聞きたい人、全員に直接意見を聞いて回るか、アンケートに答えてもらうという方法もありますが、何万という人間相手では現実的とは言えません。それに代わる方法として発展したのが、「無作為抽出（ランダム・サンプリング）」で、これは対象者全員にある番号をふって、それぞれが同じ確率で選り出されるように番号を選り出し、全員を調べる代わりとするものです。いわば対象者全員の「ミニチュア」として、代表者を選りだすわけです。選り出すための番号はコンピュータで乱数を作って決めることが多く、こうして選り出された人たちを「無作為抽出標本」と呼びます。この人たちをきちんと調べることによって、全体の姿を推し量るわけです。

こういった無作為抽出という、手間をかけず、手に入れやすい限られた情報だけに頼って（grab sample、“ひっつかみ”標本からの情報と言われます）、間違った一般論を広げてしまうことは「過度の一般化（over-generalization）」と呼ばれます。疫学ではデータを解釈する時の基本的な注意点としています。社会のオピニオン・リーダーを自認する人は、なぜ自分の意見が社会を代表できるのか、本当に代表しているのか常に反省が必要になるでしょう。意識的にか無意識にか、自分に都合の良い情報しか耳に入らなくなっているような「代表者」も時々居るようですから…。

Ⅳ. 「代表性のあるサンプリング」を行っても…?

リテラシー・ダイジェスト社は「無作為抽出」の方法を取らず、ギャラップ社はそれに近い方法を取ったことが勝因でした。しかしここでもう一つ考えなければいけないのが、リテラシー・ダイジェスト社の調査の回答率が全体の2割ほどだったという問題です。仮に無作為抽出を行っていたとしても、これほど回答率が低ければ、結果は意識の高い一部の人の意見に限られてしまい、一般化するには危険です。回答率を上げる方法として、謝礼の提供、回答しやすい調査票の工夫などいろいろ方法がありますが、努力をしても昔に比べて協力を得にくくなっているのは事実です。テーマにもよりますが、回答率は60%あれば上出来、50%台で何とか、それ以下だとキビシイかも…というのが「相場」でしょうか。いかにして対象とする人々に調査に協力してもらえるか。それは調査・研究に携わる者にとって永遠の課題と言えるのかもしれませんが。