

1936年米国大統領選挙予測－日本における標本規模の誤解

鈴木 督久 (個人会員)

1. 目的

1936年の米国大統領選挙における予測調査は、“straw-vote”から“scientific approach”への歴史的転換点として有名だが、日本では「重要な誤解」が定着したまま、恰好の教訓として流布されてきた。重要な誤解とは標本規模 (sample size) のことである。ここで文献・資料等から標本規模に関する情報を整理し、認識を再確認したい。

2. 用語

“straw-vote”あるいは“straw-poll”という用語は、麦藁を投票用具として使っていた、あるいは麦藁帽子を抱えてその中に投票してもらった昔の習慣を語源とする解説がある。麦藁を空中に投げて風向きをみる行為を、選挙における投票動向を探る風見の比喻とする説明もあるが、麦藁では重いと思われるので風見起源説はやや疑わしい。

米国の調査史を解説した Converse (1987) には“straw-vote journalism”という用語もある。雑誌・新聞などのマスメディアが世論を把握するために、多くの人が集まる場所に出かけて取材する方法として説明されている。日本語の「駅前調査」である。

“straw-vote”の定訳はないと思われるが、模擬投票・人気投票・非公式投票などの意識が候補となるだろう。公式投票の代替ではあったが、集めやすい人々や、応じた人だけを集めていた。“straw-vote”という用語には、調査の初源・調査以前の意見収集法——のような連想が含まれている。

“scientific approach”は、古い“Straw Vote Era”に対する“Modern Polling Era”として対置されているが、“scientific”は括弧付きの「科学的」であり注意が必要である。“straw-vote”の代表格は The Literary Digest であった。これにとって替わった「現代的な」世論調査の旗頭が George Gallup であった。Gallup は市場調査の経験で得た知見から、母集団を代表すると考えた属性分布から構成した層への割当標本を選挙予測にも採用した。この代表性を Gallup は「科学的」だと強調していた。統計学的な標本誤差にも言及していたが、Gallup の調査は確率標本ではなかった。層の作り方は工夫していたが、層内の有権者に関しては非確率的な有意標本であった。

3. 歴史の記憶

「模擬投票の時代から、世論調査の時代へ」の歴史的転換は 1936 年の「Digest の失敗と Gallup の成功」に象徴される。しかし割当標本で Roosevelt 当選の予測に成功した調査会社

は Gallup だけでなく何社もあったのに、なぜ Gallup だけが突出して名を残したのか。1935 年に調査会社を設立したばかりの Gallup は、Roper や Crossley よりも宣伝に熱心だった。米国には多くの地方紙があった。Gallup は 78 もの新聞社に調査レポートを売って収入を得ていた。Digest のようにメディアが調査して記事を掲載する自社モデルから、調査会社の予測結果をメディアが掲載するシンジケート・モデルへの転換であった。

新興の調査機関が自社の調査手法を宣伝する場として、大統領選挙は最高の舞台である。1936 年 11 月の投票に向けて、Digest は 8 月から郵送調査を開始。9/5 号 (24,689 票) から 10/31 号 (2,376,523 票) まで、回収票の累積集計を The Literary Digest (隔週刊で 9 号分) に掲載した。Gallup はまだ Digest が調査も始めていない 7 月の段階でコラムを寄稿し、それを各紙が編集のうえ掲載した。Gallup は選挙予測の権威であった Digest に照準を合わせ、以下のように Digest の予測得票率を、予言者からの宣告のように予測した。

If The Literary Digest were conducting a poll at the present time ... the actual figures would be in the neighborhood of 44 per cent for Roosevelt and 56 per cent for Landon.
(The New York Times. 1936, Jul. 19)

Digest の最終予測得票率は Roosevelt 43%、Landon 57% だった (二者合計を分母とした割合) ので、Gallup による「Digest 予測」の予測的中し、選挙結果は Digest の予測が失敗 (Gallup は成功) した。人々は Gallup の「予測の予言」に非常に驚いた。Gallup は Digest と同じ名簿から標本を抽出して郵送したと述べ、標本規模は Gallup(1972a)で 3,000 人だと明記している。

4. Gallup の予測調査

「Digest の予測結果を事前に予測した」3,000 人調査ではなく、Gallup 自身による本来の大統領選挙予測は、7 月時点では (1) 今回は接戦である (2) 得票率では Roosevelt が上回っているが獲得選挙人は Landon が上回っている——とコラムに書いた。どちらが勝っても傷つかない予測である。さすがに 7 月には選挙結果は断言できないが、Digest の結果は見えていた。なお、Gallup は Roosevelt の人気 (支持率) について、早くも 1934 年 2 月から月次調査を継続しており、情勢判断に有益な知見を蓄積していた。

選挙結果は、接戦ではなく歴史的な大差となった。したがって得票率と獲得選挙人の逆転もなかった。Roosevelt の獲得選挙人は 523 となり、Landon はわずか 8 (2 州のみ) という圧勝に終わった。

Gallup の調査方法は郵送と面接との併用であった。標本規模は、Gallup & Rae (1940)で、下記のように 1 回の調査で 2.5 万~4 万人であることが報告されている。

The Institute's successive samples in 1936 averaged from 25,000 to 40,000 replies, and in

states where the campaign promised to be close, the proportion of ballots was increased. Most of the Institute's ballots were collected by mail in 1936, augmented by direct interviews from groups which failed to respond to mailed ballots.

Gallup & Rae (1940)の出版に先立つ、Robinson (1937)にも同様の記述があり、数字も一致している。

On the matter of adequacy, Gallup found that his returns showed a high degree of stabilization with a national send-out of 100,000 ballots. Taking the return from both mail and personal interview canvases, each complete sample would therefore be based on from 25,000 to 40,000 replies. During the 1936 Presidential campaign, the send-out was materially increased in states in which the political race was thought to be close, the final reports of the American Institute of Public Opinion being based on approximately 125,000 replies.

当時の文献である Robinson (1937) と Gallup & Rae (1940)から、Gallup が複数回の調査を継続的に実施していたことは確実である。しかし、標準的な統計学のテキストである Freedman, et al. (2009)の記述には曖昧な部分が2点あり、誤解を生みそうである。

In fact, George Gallup was just setting up his survey organization. Using his own methods, he drew a sample of 3000 people and predicted what the Digest predictions were going to be --- well in advance of their publication --- with an error of only one percentage point. Using another sample of about 50,000 people, he correctly the Roosevelt victory, although his prediction of Roosevelt's share of the vote was off by quite a bit.

第一に、「独自の抽出法」が Digest の結果を予測する 3,000 人調査に適用されたと読める。これは Digest と同じ抽出法であるべきで、そのように実施されたからこそ Digest の結果を予測できたのである。Gallup も大統領選予測で採用した割当法(独自の抽出法)を 3,000 人調査に適用したとは述べていないし、そうすると得票率の予測値が変化するとさえ述べている。Gallup の割当法は、ここで示された 5 万人調査にのみ適用された。実は初版(1978)にはこの記述 (Using his own methods) はなかった。4 版の加筆は残念である。

第二に、Gallup が 1 回だけの 5 万人調査で予測したと読める。これは後述する日本の誤解と通じる。簡潔に省略した説明だが、読者に誤解を与えかねない。

1936 年の大統領選挙では、7 月から 10 月の 3 か月間で 11 回の調査を実施したことを、Gallup(1972b)を確認した西平 (2009) が報告している。そうすると、1936 年の選挙予測調査における標本規模は全体で 30 万人前後だと見込まれるが、各調査の正確な標本規模 (計画標本と回収標本の人数) は報告されていない。この曖昧さは Gallup の責任だが、割当法

の特性でもある。

Robinson (1937)により Gallup の郵送調査の回収率が 25%~40%であったことが分かる。これは督促や再送をしない場合の、日本における郵送調査の回収率と同程度であり納得できる。しかし 11 回の調査概要は不十分であり、毎回独立の標本で実施していないのはいか、という疑問さえ生じる余地がある。

Robinson (1937)が示した「最終報告では約 12.5 万人の回答者数」も、最終回だけ独立の大規模調査を実施したと理解してよいだろうか。あるいは 11 回の調査の総回答者数が 12.5 万人と述べているとも解釈できるが、2 万人調査を 10 回実施しても 20 万人程度には達するので、総数 12.5 万人を 11 回に分けて使ったとは考えにくい。

結局、かろうじて確からしいと思われる調査概要は、

- (1) 選挙期間中に 11 回の予測公表をした。調査も 10 回程度は実施しただろう。
- (2) 1 回あたりの標本規模は平均して 2.5 万から 4 万。5 万以上もあったようだ。
- (3) 郵送法と面接法を併用して、Gallup 方式の割当標本を構成した。
- (4) 接戦の程度で州別の標本規模は異なる。序盤から終盤への時期でも異なる。

——というところではないだろうか。

5. 日本における伝説

1936 年の米大統領選挙の情勢報道として、朝日新聞の特派員が最低でも 3 回 (10/23, 10/26, 11/3) は伝えている。現地では Gallup のリリース資料も出たと思われるし、米国各紙も報道している。10/26 と 11/3 の紙面では、Gallup の標本規模が 27.5 万人であると伝えている。おそらくその時点までの回答者の合計だろう。

戦後初の世論調査の成書だと思われる小山 (1946) は、Gallup の標本規模は 30 万人以下と書き、朝日新聞とほぼ対応している。小山は Gallup の既刊書等を戦前に読んでいた。

戦後 2 冊目の成書だと思われる吉田・西平 (1956) の記述は曖昧で、Digest の 1%未満としている。約 2 万人を超えない、ということになるが人数は示されていない。

戦後 3 冊目以降の出版に関しては、完全な追跡はできていない。特に 1960 年代の書籍は未見だが、佐藤他 (1976) は Gallup の予測調査の標本規模が 3,000 人だと明記した。そこから現在に至るまで 10 冊以上の成書が 3,000 人と書いている (2,000 人、1,500 人と書いた成書もある)。

1936 年の教訓は「模擬投票」より「標本調査」、つまり標本の代表性の重要性を認識した点にある。しかし日本では、偏った「大量観察」よりも代表性のある「小規模標本」——という側面も記憶に強く残った。たとえば、佐藤他 (1976) の記述を見よう。

ところが 1936 年の選挙において、約 200 万人から得た調査結果により、57%の得票率でランドンの当選を予測したが、選挙の結果は 43%の得票率で落選を予測したルーズヴェルトが当選するという惨めな結果となった。これをきっかけに同社の社運は次第に傾

き、ついに廃刊の憂き目を招いたといわれる。

これに対し、ギャラップ (George Gallup) は、わずか 3,000 人について調査を行なったにもかかわらず、投票結果の予想をみごとに的中させた (表-3)。この結果、ギャラップは、その後リテラリー・ダイジェストに代わって調査界に君臨することとなった。

200 万人に対して「わずか」3,000 人！——これは強い衝撃を受ける。1970 年代以降の数十年にわたって日本で 3,000 人が定着した理由であろう。いつから 3,000 人という誤りが定着したのか不明確だが、誤りが引き継がれた原因は、強烈な印象が残るからだろう。これまで、3,000 人という標本規模に疑問を呈した議論は、杉野 (2006) しか見当たらない。なぜ 3,000 人という数字で誤ったのかも原因不明だが、Digest の予測値を予測する 3000 人調査と混同した可能性が高い。Gallup が別の文脈でもしばしば 3,000 人という標本規模の適切性を主張していたことも影響した可能性がある。

1936 年の米国の有権者数と選挙区数は、日本における現在の参院選と同程度である。報道機関が実施している参院選の情勢調査の標本規模は数万人である。選挙制度が異なり、大統領選では全選挙区で候補者が同じとはいえ、選挙区別の調査は必要なので 1 回の調査は数万人規模になるだろう。しかも選挙期間の長い大統領選では複数回の調査を実施する。1 回の 3,000 人調査で十分だという議論は、全国世論調査なら成立するが、大統領選の予測調査である。調査専門家として疑問を持つべきであったろう。

謝辞

吉野諒三氏、松本渉氏、福田昌史氏から文献の教示、討論での示唆を頂きました。

文献

Converse, J. M. (1987). *Survey Research in the United States: Roots and Emergence 1890-1960*, The University of California Press.

Freedman, D. , Pisani, R. , Purves, R. (1978). *Statistics*. New York, Norton.

Freedman, D. , Pisani, R. , Purves, R. (2009). *Statistics, Fourth Edition*. Viva Books.

Gallup, G. H. & Rae, S. F. (1940). *The pulses of Democracy. the public-opinion poll and how it works*. Simon and Schuster.

Gallup, G. H. (1972a). *Sophisticated Poll Watcher's Guide*. Princeton Opinion Press.

Gallup, G. H. (1972b) *The Gallup Poll, Public Opinion 1935-1971*. New York, Random House.

Robinson, C. E. (1932). *Straw Votes: A Study of Political Prediction*. New York: Columbia University Press.

Robinson, C. E. (1937). *Recent Developments in the Straw-Poll Field*. *The Public Opinion Quarterly*. Vol. 1, No. 3.

- 小野裕亮 (2021). 私信. <https://suzuki-tokuhisa.com/poll>
- 小山栄三 (1946) 輿論調査概要. 時事通信社.
- 佐藤彰・鈴木栄・船津好明 (1976) 世論調査－設計と技法－. 技興社.
- 杉野勇 (2005). 1936 年大統領選の実際－Literary Digest と Gallup 再訪－. 「関連社会科学」第 15 号.
- 鈴木督久 (2021). 世論調査の真実. 日本経済新聞出版.
- 西平重喜 (2009). 世論をさがし求めて－陶片追放から選挙予測まで－. ミネルヴァ書房.
- 吉田洋一・西平重喜 (1956) 世論調査. 岩波新書.