

調査についてのリテラシー向上のために



NHK放送文化研究所世論調査部 たかはし こういち
高橋 幸市 氏

1. はじめに

「NHKの世論調査」というテーマで講演の依頼をいただきましたが、昨年のセミナーで朝日新聞社の渡辺さんが、朝日新聞社の世論調査について幅広くご紹介していらっしゃいます。NHKについても朝日新聞社と大きく変わるところはありませんし、労働調査協議会の機関紙に渡辺さんの講演内容が掲載されていますので、そちらをご覧くださいと思います。

それでは、私は本日、何をお話しすべきかと考えました。今日、このセミナーに参加されている方のうち、どのくらいの割合の方がアンケートや世論調査を実際におやりになるかは分かりませんが、恐らく、ご自分で調査をされるよりは、調査結果をご覧になるほうが多いと勝手に考えまして、調査を実施する時ではなく、調査結果を利用される際に注意すべき点をいくつかお話しして、世論調査に関するリテラシー向上に少しでも役立ていただければ、と思います。

そして、本日の私の話の要点は、調査のやり方によって利用の仕方が異なるのですが、私の目からは、調査の仕方が異なるにも関わらず同じように評価しているなど、少し混乱しているように思えるのです。そこで、本日は、調査の仕方によって利用の仕方が異なるという点についてお話ししたいと思います。

2. 世論調査とは

まず、世論調査とは何を指すのかを考えてみたいと思います。「リアルタイム世論調査.net」といホームページがあります。運営は世論調査やマーケティング・リサーチを受託している会社が行っております。ちなみに、この会社は日本の大手の世論調査会社のほとんどが加盟している「日本世論調査協会」に加盟しているとホームページには書いてあります。

このホームページでは、アクセスした人は好きな質問を選んで自由に回答できることはもちろんのこと、質問も自由に作成して掲載できるようです。インターネットで検索しますと、この例に限らず、会社や個人が運営している、同じような「世論調査」が数多くあります。

そこで、皆さんは、このような調査を「世論調査」と考えますでしょうか。

「調査」については、この場に出席されている方の中では、イメージが共通しているかと思いますが、「世論」については、皆さんの中でもかなり幅広いイメージを持たれていると思います。不特定多数の人から集めた意見の集合、を「世論」という人もいます。この場で、「世論」の定義についてお話ししていたのでは、時間が足りませんし、私の能力を超えるテーマでもありますので、専門家の間でも「世論」の定義はいろいろある、とだけ申し上げておきます。したがって、「世論」の定義がいろいろありますので、「世論調査」を一律には定義することができないのです。

逆に言えば、調査結果をご覧になる際には、「世論」をどう定義して行った調査なのかを吟味する必要があります。

私が承知していて、そして、世論調査を行う者にとって、一番分かりやすい定義は、日本の世論調査の発展に多大な貢献をなされた林知己夫さんの定義です。それは、「世論は、世論調査の方法によって明らかにされたものであり、調査の対象の集まりの、ある時点における、ある質問群に対する意見分布、構造、あるいはその動態を言う」というものです。

先ほどの「不特定多数の意見」とほとんど変わらない、と感じられる方も多いと思います。実は、この定義だけでは林先生の真意は分からないのです。林先生は統計数理研究所の方でありまして、サンプル調査の理論など、多くの研究をされましたが、その理論の検証のためもありまして、世論調査をご自分で実施されてきた方です。従いまして、先ほどの定義の前に、「サンプルが正しく抽出されている」という前提が付きまします。

「サンプルを正しく抽出する」ためには、どうしても「母集団」という概念について知っておく必要があります。そこで、サンプルの抽出についてお話ししますが、その前に、母集団についてお話しします。母集団について理解していないと、サンプルについて分かりませんので。

3 . 母集団について

図1をご覧ください。この図は調査について簡単に示したものです。そもそも調査を行う目的は、ある集団について何らかの情報を得たい訳ですね。知りたい集団のことを「母集団」と言います。図では真ん中に大きな円で示してあります。「母集団」とは、例えば「日本人」とか、あるいは「東京都に住んでいる人」、さらには「ある企業に勤めている人」とか、「ある企業の労働組合加入者」を指します。

「母集団」という言葉は慣れていないと難しいと思いますが、実はこれを明確にしておかないと、何について調査したのかが分からないこととなります。「労働組合加入者」について調べたいとする場合は、基準がはっきりしていますし、名簿もありますから、「母集団」は明確なのですが、必ずしも「母集団」が明確でない調査もあります。冒頭ご紹介した「リアルタイム世論調査.net」では、「母集団」は何になるのでしょうか？果たしてすぐに「誰を対象とした調査なのか」が明確に分かるでしょうか。

調査を行う場合には、まず、対象、つまり母集団を明確にする必要があります。これは、全員を調べる調査でも同様です。全員とはどこまでの範囲の人を指すのかを明確にしておかないと、調査用紙を配付しようにも配付できませんよね。ですから調査を行う場合には、まず「母集団」、つまり、「どこまでの範囲の人を対象とするか」を明確にしなければなりません。

4 . 全数調査にも誤差がある

次に、サンプルについてお話ししますが、最初に、全員を調査する全数調査についてお話ししたあとで、サンプルについてご説明します。

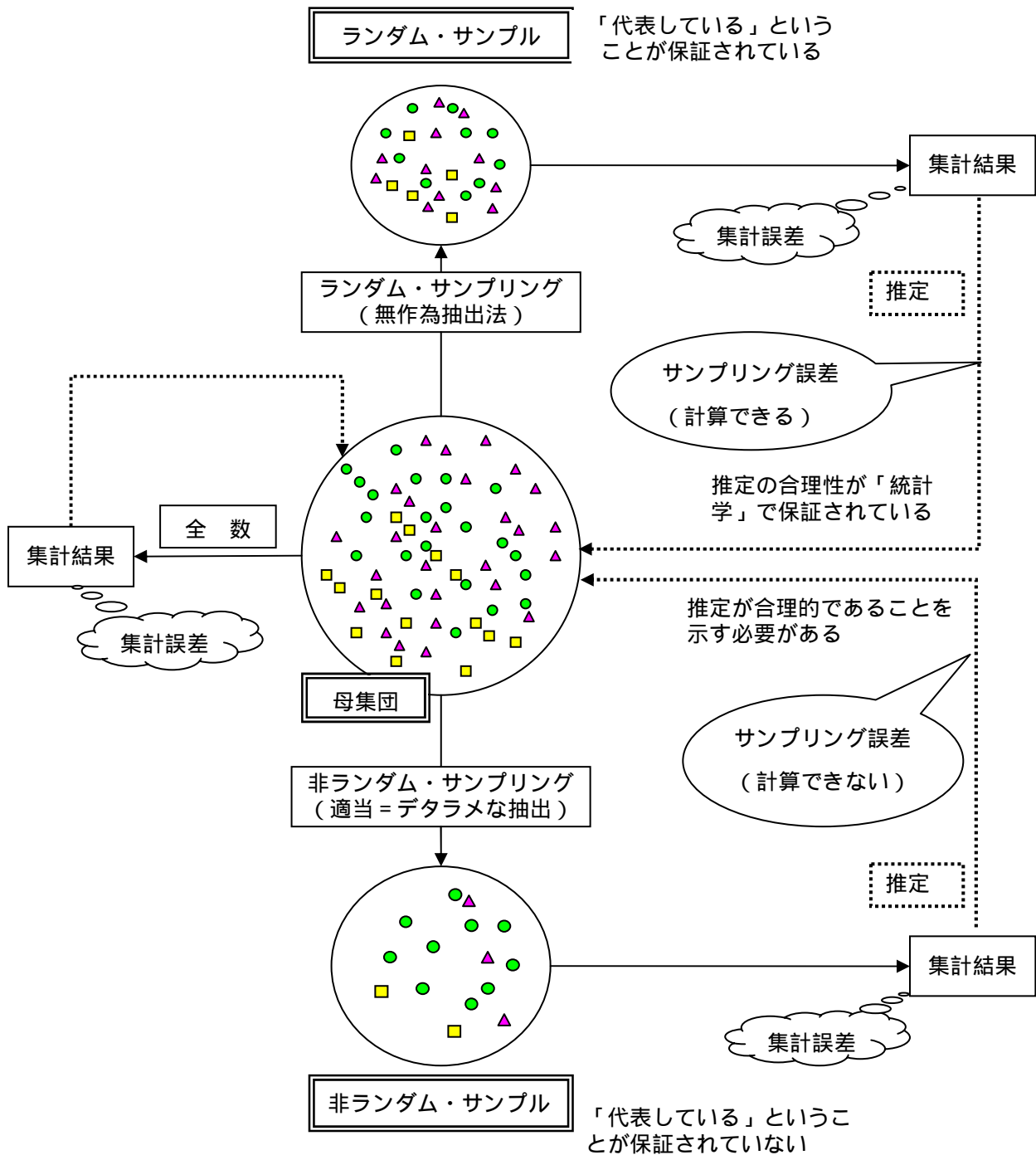
では、図1の左半分をご覧ください。母集団の構成員全員について調査する「全数」調査が書いてあります。構成員全員についてきちんと調査すれば、母集団について正確な情報が得られますから、安心して調査結果を利用できるでしょう。

ただし、私が、世論調査に関する議論が混乱していると申し上げている一つの原因に「誤差」に関する点があります。そこで、図1にも書いておきましたが、全数調査にも誤差はあります。質問に対する回答を記録し、パソコンに入力して%を計算するなどの方法をとる場合には、回答の記録ミス、パソコンへの入力ミス、など様々なミスがある、と想定しなければなりません。この誤差は、

集計だけに関わらないものであり、専門的には、これから説明します「サンプリング誤差」以外の誤差を指していますが、図1では分かりやすいように「集計誤差」と表示してあります。

そして、この集計誤差はどんな調査でも発生します。ただし、この誤差はきちんと調査し、集計していれば、あまり大きくはないので無視することが多いです。

図1 調査の種類（サンプリングの方法から見て、概要）



あるいは、調査結果は数字で示すことが多いのですが、この誤差は数字で示すことができません。したがって、この誤差について考慮する場合には、回答結果を%で示す時に小数点以下を四捨五入したりして整数で示すなど、あまり細かい数字は評価しないという配慮をすることもあります。

無視するにしても、配慮するにしても、この誤差はどんな調査にも発生する誤差であり、また、「サンプリング誤差」とは異なるものである、と認識しておくことが大切です。

5 . ランダム・サンプリング

続いて、サンプル調査について説明致します。

ただし、最初に申し上げておきますが、サンプルを選ぶ方法の中で、どの方法が一番優れているかということは決められません。調査の目的、別な言い方をすれば、調査結果をどのように利用するかによって、最適な選び方が決まるのです。冒頭申し上げた混乱の原因の二つ目を、私はこの点にあると考えています。どれが一番優れた方法かを決められないのに、決まっているかのような議論をすること、あるいは、サンプルの選び方によって調査結果の利用の仕方も異なる、ということが明確に認識されていないことが、混乱の原因になっていると考え、これからお話しする内容を決めたのであります。

では、図1の説明に戻ります。全員を調べるのは経費がかかる、あるいは時間がかかって大変だという場合には、一部の人のみについて調べます。そうすると「サンプルの選び方」ということが問題になってきます。そして、サンプルの選び方によって、その調査結果の利用の仕方が異なります。ですから、経費や時間に問題がないのであれば、全員を調べてしまったほうが、結果的には効率的な場合があります。

そのサンプルの選び方について、図1では右側の上半分と下半分に分けて書いてあります。まず、

上半分について説明しますが、上半分は「ランダム・サンプリング」についての説明であります。ランダム・サンプリングという言葉も聞いたことがあると思います。日本語では「無作為抽出法」と言われます。ランダムにしても、無作為にしても、分かるようで分からないという人がいらっしゃるかも知れません。

「ランダム・サンプリング」について説明した本などでは「乱数表を用いた抽出」などと書かれていることが多いのですが、正しい乱数表を手に入れるのは実は大変ですし、それを正しく使うのも大変ですので、多くの場合、「系統抽出（等間隔抽出）」という方法を用います。具体的には「スタート番号」と「インターバル」を用いる方法です。

例えば、1,000人の従業員から100人を選ぶとします。その場合には1,000を100で割ってインターバルを求めますので10となります。そして10より小さい数字一つを乱数表から選びます。10より小さい数字を選ぶのは、そうしないと100人抽出できないからです。

ただし、この方法の場合、注意しなければならない点があります。それは、抽出に使用する名簿に規則性がある場合、その周期とインターバルが同じになっていないかという点です。例えば、職場のグループが10人単位で構成されていて、その順に名簿が整理されていると、同じような立場の人が選ばれてしまいます。名簿の整理の間隔とインターバルが同期していないか必ずチェックする必要があります。この点を注意すれば、系統抽出（等間隔抽出）は非常に効率的な方法です。

ここでのキーワードは「ランダム」です。最近はいろいろな抽出方法がありますが、要するにランダムになっているかいないかが重要なのです。

また、ランダム・サンプリングの説明として「等確率抽出」という言葉も良くみかけます。例えば、先ほどの1,000人の従業員から100人を選ぶ際に、1,000人の従業員の選ばれる確率が等しい、

ということです。この場合、使用する乱数表は4桁乱数表ですが、正しい乱数表は4桁の数字が表われる回数がすべて等しくなっていますので、乱数表を用いた抽出方法は等確率抽出法となります。

ただし、もう少し詳しく申し上げますと、抽出時の確率が等しくなくても、結果的に等しければ良いのです。結果的に等確率になっていれば良いのです。そのためには、一つ一つのサンプルの抽出確率が分かっている必要はありませんが、分かっているならば、集計する時に抽出確率に応じたウェイト集計を行えば、ランダム・サンプリングをしたことと同じになります。ごく稀に、このような方式が採用されることもありますので、念のために申し上げます。

6. 非ランダム・サンプリング

では、次にランダムでないサンプリングについて説明します、図1の右側の下半分についてです。つまり、適当、あるいはデタラメにサンプルを選ぶ方法についてであります。ここでは一括して「非ランダム・サンプリング」としてあります。

ところで、デタラメについて説明するのは結構難しいことです。そこで、サンプル調査の歴史を辿って、その説明に替えたいと思います。また、これからアメリカの例をいくつか紹介しますが、日本と違ってアメリカでは住民基本台帳のような国民全員が登録されている名簿はありません。そのために、これから紹介するような様々な方法が試みられてきました。

(1) リテラリー・ダイジェスト社の教訓

～数を集めてもダメ

まず、最初の例はサンプルをたくさん集めてもダメだということです。表1をご覧ください。これは1936年のアメリカの大統領選挙についての調査の結果ですが、実際には60%の得票率でルーズベルトが当選しました。ところが、リテラリー・ダイジェスト社は落選したランドンが得票率57%で当選すると予想してしまいました。

古い話ですので数字については、諸説があります。ここに書いてある数字と、これから紹介する数字はだいたいの概数だと考えてください。

表1 1936年大統領選挙

(%)	ルーズベルト	ランドン	計
実際の得票率	60	40	100
リテラリー・ダイジェスト	43	57	100
ギャラップ (Gallup)	56	44	100

正確な数字については、諸説がある

鈴木督久 (<http://suzuki-tokuhisa.com/ushigome/gallup.html>)

西平重喜『世論をさがし求めて』(2009年、ミネルヴァ書房)

リテラリー・ダイジェスト社

1,000万人に模擬投票用紙を郵送し、約237万6,000人から回答を得る

使用名簿：電話加入者名簿、自動車所有者名簿、愛読者名簿

ギャラップ

割当法 (quota method) により抽出した有権者3,000人から、面接法で回答を得る

quota method：割当法 (地域、性、年齢、社会階層 = 職業・収入、人種)

この時のリテラリー・ダイジェスト社は電話加入者と自動車所有者、さらにはリテラリー・ダイジェスト誌の愛読者、合わせて約1,000万人に郵便を送り、そのうちの230万人を超える人から回答を得ました。一方、ギャラップは割当法という方法で、230万人に比べればわずか3,000人からの回答で、みごとにルーズベルトの当選を予想したのです。リテラリー・ダイジェスト社は、それまでも同じ方法で当選者を当てていましたので、なぜ、この時は外したのかが問題になり、検証した結果、調査の対象としたリテラリー・ダイジェスト誌の愛読者や電話加入者、自動車所有者は有権者を代表していなかったということが判明したのです。つまり、これらの人は富裕層、豊かな人です。しかし、この時の選挙で投票した人は4,000万を超えていたと言われますが、その中には、雑誌を買ったり、電話や自動車を持つことができない人もたくさん含まれていたのです。この失敗からの教訓は、いくらサンプルをたくさん集めても目的に合っていなければダメだということです。

リテラリー・ダイジェスト社の目的は、有権者が誰を大統領に選ぶかということをはっきりさせることにありました。ですから、母集団は「実際に投票した人」になります。しかし、投票が終了しないと「投票した人」は分かりませんので、事前に予測するためには、「投票に行く可能性が高い人」と読み替えるしかありません。その上で、貧富の差、つまり、貧しいとか豊かとか関係なく選ばなければならなかったものを、豊かな人たちだけを選んでしまった点に、失敗の原因があったのです。同じ方法で、それまではリテラリー・ダイジェスト社は当選者を当てていました。この時の選挙で、ルーズベルトは有名なニューディール政策を掲げて戦ったのですが、この政策は豊かでない人びとの支持を得ました。そうした人びとをサンプルに選ばなかったことが失敗の原因だったのです。

なお、1,000万の人、一説には2,000万人という

数字もありますが、とくかく何故こんなにたくさんの人に郵便を送ることができたのかという疑問が生じます。当時でも大変な経費ですよ。そのカラクリは、どうも、雑誌の販売拡大を兼ねていたようです。つまり、リテラリー・ダイジェスト誌はこんな雑誌で、たいへん面白く役に立ちますから、お買い求めください。そして、よろしければ、今回の大統領選挙で誰に投票するつもりかを、同封したはがきでお返しく下さい、というような方法だったようです。

一方、ギャラップのほうは3,000人しか調べていないのに、見事的中しました。この「3,000人」という数字もいくつかの説がありますし、また、割当法とはどんな方法かもまだ説明していませんが、この点に関しては、次のパートで触れます。

(2) ギャラップ (Gallup) の教訓

～割当法 (quota method) も代表していない

表2をご覧ください。これは1948年のアメリカ大統領選挙に関する調査の結果です。

この時の選挙では、トルーマンが約50%の得票を得て当選していますが、ギャラップを含め、当時の主だった調査機関の予測がすべて外れました。

この表は、先ほど紹介しました林知己夫さんと同じく、統計数理研究所で世論調査を中心としたサンプル調査や解析方法などを長い間研究し、日本の世論調査の発展に大きな貢献をされている西平重喜さんの『世論をさがし求めて』という本から引用しましたが、一部が違っていると思われるので、私が勝手に修正してあります。

ギャラップをはじめ、各社とも割当法というサンプリング方式を用いていたようですが、ギャラップの方式が一番細かく指定していたようです。そこで、先ほど紹介した西平さんの本(p76)から、ギャラップの具体的な方式について引用しますと、「市町村の人口階層別(4段階)、性別、人種(白人・非白人)、階層(職業、4段階)が使

われていた」とあります。つまり、国勢調査の結果から性別や、白人と非白人の割合などが分かりますから、市町村ごとに、これらの割合が国勢調査の結果と一致するようにサンプルを選んで調査するというものです。指標は4つありますが、人

口階層は市町村ごとに調査地点を割り当てますので、個人でみた場合には3つの指標となり、区分を掛け合わせますと2×2×4で16になります。この区分ごとに何人を調査するかを決定します。

表2 1948年大統領選挙

(%)	デューイ	トルーマン	サーモンド	ウォーレス	計
実際の得票率	45.1	49.5	2.4	2.2	99.4
ギャラップ	49.5	44.5	2.0	4.0	100.0
クロスレー	49.9	44.8	1.6	3.3	99.6
ローパー	52.2	37.1	5.2	4.3	98.8

西平重喜『世論をさがし求めて』(2009年、ミネルヴァ書房)を元に修正

市町村よりも狭い地域について、国勢調査の結果が公表されているかどうかは分かりませんが、想像するに、市町村の中でいくつかの地点を調査しますので、それを全部合わせた結果が市町村単位で国勢調査の結果と一致するように、サンプルの人数を決めたのだと思います。

なお、西平さんの本では、調査の方法については明確には書いてありませんが、郵送法ではなく面接法であったと思われます。ですから、調査員は指定されて割当表に基づいて、例えば、男性の白人で職業階層が「上」の人を3人、女性で、しかも白人でない人で、職業階層が「下から2番目」の人を2人、という風に面接し、合計で15人を調査しなさいと指示される訳です。なお、この時には年齢は指標の中には入っていなかったようです。

先ほどリテラリー・ダイジェスト社が失敗し、ギャラップは成功した例として紹介しました1936年の大統領選挙でも、ギャラップは同じ割当法でサンプリングしています。

割当法の考え方は、どの大統領に投票するかは、割り当ての指標にした性(男か女か)と人種、そして職業階層によって決まっているというものです。それまでの調査の経験からアメリカでは、こ

うした割当をすれば、選挙の結果を予測できると考えていたのです。しかし、1948年は失敗しました。

したがって、この時の教訓は、例えば性とか、年齢とかを母集団の構成比に一致するようにサンプリングしても、その他の指標は母集団と一致するとは限らないということです。

その背景には、簡単に言えば次のようなことがあります。割当表に該当する人がいたとしても、その人に断られたら、次の人を探さなければなりません。そうすると、簡単に予想できると思いますが、調査に協力してくれる人だけがサンプルに含まれることになります。ところが、サンプルに含まれない、つまり、調査に協力しなかった人も投票しますので、実際の結果とはずれてしまうことになった訳です。

割当法というのは、現在の日本でも結構行われています。NHKでも使用しています。アメリカでは今日でも指標は先ほど紹介したように、性別と年齢以外に、人種とか宗教とかを用いていますが、日本では性別と年齢だけが多いようです。割当法でサンプリングした調査の結果は、たとえ性や年齢が母集団の構成比に一致していたとしても、その他の指標は一致するとは限りません。つまり、

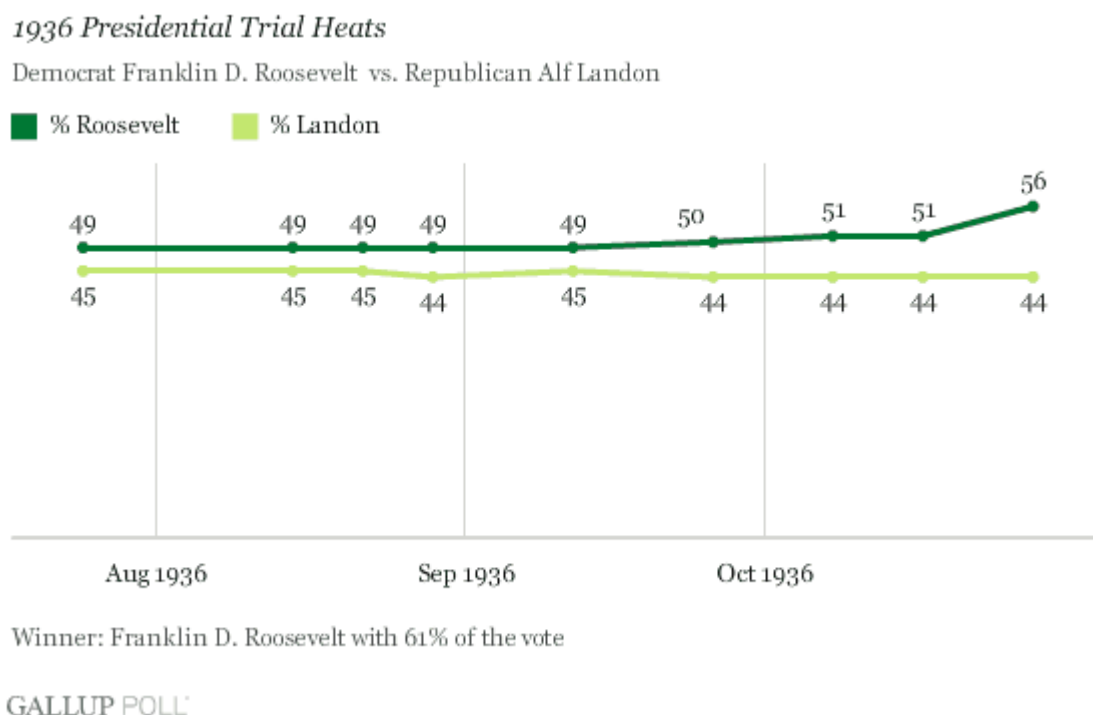
調査を行う者がほしいと思っているもの、例えば投票予定だとか、自動車の好みだとかの割合は母集団の割合に一致するとは限りませんので、相当慎重に利用する必要があります。

では、図2をご覧ください。これは、ギャラップのホームページから切り取ってきたものです。ギャラップでは1936年以降、昨年までの大統領選挙に関する調査の結果をホームページに掲載しています。図2には、先ほど紹介しました1936年と1948年の結果を示してあります。この図を見ますと、大統領選挙に際しては何回か調査を行っているようでして、私が紹介しました数字は、グラフのそれぞれ右端の結果です。最後の調査、つまり、最も投票日に近い調査の結果ということになります。ただし、これらの調査がどのようなサンプルを対象に、何人から調査した結果なのか、な

どが表示されていません。いろいろな資料を見ますと、サンプルから得られた結果そのものでなく、修正しているようです。その修正も、アメリカでは通常行われる「サンプル構成の歪み」を補正する方法だけでなく、経験に基づいて、例えば民主党の候補は2%上乘せする、などの補正をしているようですが、はっきりしたことは分かりません。ですから、調査の結果というより「調査の結果に基づく、大統領選挙に関する予測」と考えたほうがよいかもしれません。

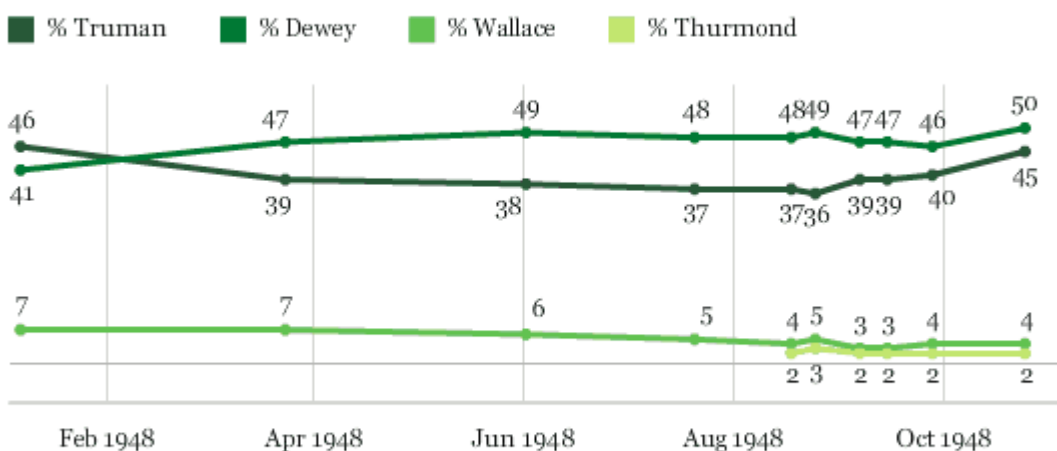
しかし、それほど大きく修正している訳ではありませんので、私の申し上げたい主旨は変わりません。割当法でサンプリングして得られた調査結果を、単純に母集団の分布と一致すると考えることは危険であるということです。

図2 ギャラップのHPから



1948 Presidential Trial Heats

Democrat Harry Truman vs. Republican Thomas Dewey
vs. Progressive Henry Wallace vs. "Dixiecrat" Strom Thurmond



Winner: Harry Truman with 50% of the vote

GALLUP POLL

出典：<http://www.gallup.com/>

(3) 2つの失敗からの教訓

～サンプルが偏っていれば、必ず失敗する

次に3つ目の教訓です。1936年の例も、1948年の例も、それまでは成功していたものが失敗したものでした。多少ずれていても、何回も継続して調査すれば、正しい結果を得られる、ないしは傾向はつかめる、という話を聞くことができます。しかし、これも根拠がありません。これまでお話しした例では、ギャラップは1936年から1948年の失敗するまでの間、10年以上に渡って「正しい」結果を入手していたと信じていたのです。にもかかわらず、最終的には失敗しました。

7. 非ランダム・サンプル調査の活用

ランダムでないサンプルを利用して、誤った利用をすると失敗する例を紹介してきましたが、最初に申し上げましたように、調査の方法やサンプリングの優劣は、単独では決められません。調査の目的に応じて評価されるものです。

これまでは、非ランダム・サンプル調査の欠点ばかりをお話ししてきましたが、非ランダム・サンプル調査は現在でも盛んに行われていますし、利点、メリットもありますので、その活用方法についてもお話ししないと不十分なことになるでしょう。しかし、実は私は非ランダム・サンプル調査の活用方法が分からないので、ランダム・サンプル調査を活用する仕事を選んだ人間ですので、この話は得意ではないのです。本当のところ、非ランダム・サンプル調査をどう活用すれば良いか分かりません。

しかし、何とかお話ししてみましよう。非ランダム・サンプルの活用の一つに、「関係性」に注目するということがあります。関係性とは、例えば「Aという意見に賛成した人は、Bという意見にも賛成する」というものです。非ランダム・サンプル調査の結果で「Aという意見に賛成する人」が30%いたとしても、母集団では何%いるかは分かりません。しかし、「Aという意見に賛成する人で、Bという意見にも賛成する人」が1人でもいれば、母集団でもそう考える人がいること

は事実です。強調しますが、「Aという意見に賛成する人で、Bという意見にも賛成する人」が母集団で何%いるかは分かりません。少なくともそういう思考の関係性を持った人が存在することは間違いないということです。

別な活用事例は、マーケティング・リサーチ業界で良く行われているものです。新しい商品を開発したり、あるいは、商品を改善したりする場合に、非ランダム・サンプル調査の結果が利用されます。NHKでの番組開発なども同じようなことが行われております。この時、データとして入手したいのは、母集団の分布ではなく、開発や改善に役立つヒントです。自動車の色は赤が好きという人が、全国で30%いるということが分かって、商品開発などの現場ではほとんど役に立ちません。うまく説明できませんので、イメージとして捉えていただきたいのですが、「この形なら赤だけど、この形ならピンクね」と考える人がいる、ということが分かるほうが重要な場合があります。私が承知している限りですが、ランダム・サンプル調査の結果だけ、つまり、母集団の分布が分かる調査結果だけでは、ヒット商品は生まれません。商品開発などではデータだけでなく、利用する人間の判断が重要になります。判断に役立つデータであれば、何も母集団についてのデータでなくても構わないのです。

さらに、この場合、もし調査の結果を母集団の分布と同じものだと判断して使用しても、その責任は判断した人、ないしは組織にあります。母集団の分布と同じだと信じた結果、商品が売れなく

ても責任は自分でとればよいのですから、そう活用しても差し支えないと思います。単純に調査結果を公表することと、その結果を利用して政策なり、商品開発の方向なりを決定することとは意味が違うのです。非ランダム・サンプル調査の結果は、母集団には単純には当てはめられないということを確認した上で使用しているのなら、問題はないのです。

8 . 計算できる誤差、できない誤差

では、図1をもう一度ご覧ください。先ほどまでの説明で省略した部分について、お話しします。

「集計誤差」についてはすでにお話ししましたが、図1にはもう一つの誤差が表示されています。母集団全体について調査したのではなく、その一部を取り出したのですから、誤差が生じます。もう少し厳密に申し上げれば、サンプルから得られた結果を母集団に適用する際には、誤差を考慮しなければならぬ、ということになります。もともと母集団について調べたかったのですから、調査の結果を母集団に当てはめなければ意味がありません。それを統計学では「推定」と言います。そして、推定する際に生じる誤差を「サンプリング誤差」と言いますが、この誤差はサンプルの選び方によって異なり、ランダム・サンプルなら、その誤差は計算できますが、非ランダム・サンプルでは計算できません。この点が、ランダム・サンプルと非ランダム・サンプルの大きな違いです。

サンプリング誤差（単純ランダム・サンプリングの場合）

$$\text{誤差} = k \times \frac{\sqrt{N-n}}{\sqrt{N-1}} \times \frac{\sqrt{p \times (100-p)}}{\sqrt{n}}$$

$$\text{誤差} = k \times \frac{\sqrt{p \times (100-p)}}{\sqrt{n}} \quad \because \frac{\sqrt{N-n}}{\sqrt{N-1}} \approx 1$$

k : 信頼度係数、 N : 母集団の要素数、 n : サンプル数（有効数）、 p : 比率（%）

$k = 1.96$ （信頼度95%、危険率5%の時）

一口にランダム・サンプルと申し上げていますが、具体的なサンプリングの方法はいろいろあります。そして、その方法に応じて誤差の計算式は異なります。それを説明していたのでは時間が足りませんし、私の能力を超えますので、関心のある方は統計学の本をご覧くださいとして、ここでも最も簡単な誤差式を紹介します。

この式は、「単純ランダム・サンプリング」を行った際の誤差式です。他のランダム・サンプリングの式は、少し複雑になりますので、通常はこの式を覚えていて代用します。

また、「N」つまり母集団の要素数は「100」以上、「n」つまりサンプル数は「30」以上、を目安として使ってください。それより小さい場合には、別の式を使用します。

この式の中の「k」は「信頼度係数」と呼ばれます。少し面倒な話になりますが、誤差を計算する場合には、どの程度の幅で見込むかを決めなければいけません。サンプル調査から得られた結果、例えばそれが30%であったとします。その際に、母集団の値も30%だと推定することも可能です。ですが、それは危険な感じがすると思います。母集団のすべての要素、全員について調べた訳ではないので、ぴったり一致していると主張するのは不安ですよね。ですから、少し幅を持たせて推定するほうが良いと思うでしょう。その幅を持たせる時に上記の誤差式を用いれば良いということが、統計学で明らかになっています。例えば、「k」を「1.96」にすれば、母集団の値とずれる可能性は5%しかない、ということが明らかになっています。ずれる可能性を危険率と言います。その逆を信頼度と言います。薬の効果を決定する場合や、製品検査の場合などには、もっと信頼性を高める必要があります。例えば、99%の信頼度、つまり、ずれる危険を1%にしたい場合には「k」は「2.58」にしますが、社会学的な調査では、通常、信頼度は95%を使用します。

そして、誤差 と誤差 が示してありますが、

誤差 のほうは $\frac{\sqrt{N-n}}{N-1}$ がありません。その右に書いてありますように、 $\frac{N-n}{N-1}$ は、Nやnがある程度大きければ、ほぼ「1」になります。例えば、母集団に含まれる人が10万人で、そこから1,000人をサンプリングした場合には、約0.99です。なので、ほぼ「1」です。

なぜ、誤差 のような簡略化した式を紹介したかと言いますと、この式を記憶して暗算で計算するのに便利だからです。最近Excelなどで簡単に計算できますし、Nやnがどの程度大きければ $\frac{N-n}{N-1}$ を「1」と見なして良いか分からないこともあるでしょうから、誤差 を計算したほうが、いろいろ考えなくて良いので楽だということも言えます。

ところで、誤差 も誤差 も、微分が得意な方は50%の時に誤差が最大になるということが分かります。そうでない方は、Excelなどで具体的に計算して確かめてください。誤差は大きく見積もっておくほうが安心ですし、一つ一つの数値について計算するのが大変な場合には、50%の場合だけ計算しておいて、すべての数値について、その誤差を適用するというだけでも良いのです。

それでは、暗算で誤差 を計算してみましょう。

例えば、100人のサンプルから得られた調査結果が50%であったとすると、分母の「n」が100で、分子の「 $P \times (100-P)$ 」は「 $50 \times (100-50)$ 」ですから2,500になります。したがって、ルートの中は「 $2,500 \div 100$ 」ですから「25」です。そして、「ルート25」は「5」です。信頼度係数は「1.96」ですと暗算が大変ですので「2」と見なします。信頼度係数をまるめる場合には大きいほうにします。「1」のほうが暗算が楽だからといって「1.96」を「1」と見なしてしまっただけは信頼度が95%よりずっと低くなってしまいます。結局、ルートの中は「5」で、信頼度係数の「k」は「2」と見なしましたから、誤差は「10」になります。

つまり、100人のサンプルから得られた調査結

果が50%であった時、誤差を10%見込めば、母集団での値はその中に95%の確率で含まれることとなります。別の言い方をすれば、母集団での値は、 $50\% \pm 10\%$ 、つまり40%から60%の間にある可能性は95%であるということになります。

もう一つの例を示しますが、誤差は「P」が50%の時に最大になると言いました。先ほど計算しましたように、50%の時はルートの中の分子は2,500です。そして「ルート2,500」は「50」です。ですから、いちいち計算しないで「50」を記憶しておくこともできます。

あとは分母の「n」がルートの中にある時に整数になる「n」を考えれば楽です。「n」が100だったら「ルート100」は「10」です。「n」が400なら「20」ですし、900なら「30」です。暗算で計算するためには、このように計算しやすい例でやれば良いのです。なお、「n」は分母にありますから、実際のサンプル数より小さい数で計算すれば、誤差が大きくなるので安心です。例えば、1,000人のサンプルの場合で計算する際には、2乗して1,000より小さくなる数を探します。先ほど挙げた例から「900」としましょう。するとルートの中は、「 $50 \div 30$ 」ですから、ほぼ「1.7」くらいになります。それに信頼度係数の「2」を掛けて、結局、誤差は「3.4」になります。

サンプル調査の結果には、誤差が伴います。そして、ランダム・サンプルの場合には、先ほどの式で誤差が計算できます。しかし、調査結果のそれぞれの数字には誤差をつけて表示しません。そうすると、見難くなるからです。その替わりとして、誤差を計算するための情報を別に表記しておくのがルールです。その表記の仕方について、次にお話しします。

9. 調査の保証書

次の枠の中をご覧ください。これはNHKが実施した世論調査に関する情報です。NHKでは、

「1」の「調査目的」から「6」の「調査有効数、有効率」までを「調査の概要」と言っています。そして、原則としてすべての調査に「調査の概要」をつけることになっています。放送などでは、時間やテレビ画面の制約があって、すべてを紹介できない時には、一部省略することがあります。

デジタル放送調査 2009 単純集計結果

【調査の概要】

- | | |
|------------|--|
| 1 調査目的 | デジタル放送やデジタル機器の普及に伴って、放送に関連したサービスの多様化が進み、視聴者のメディア接触方法も変化しつつある。こうした中、放送を中心としたメディアがどのような形態で利用され、視聴されているのか、実態と意向を把握する。 |
| 2 調査時期 | 2009年9月5日(土)～13日(日) |
| 3 調査方法 | 配付回収法 |
| 4 調査対象 | 全国の16歳以上の国民 |
| 5 調査相手 | 住民基本台帳から層化無作為2段抽出
3,600人(12人×300地点) |
| 6 調査有効数(率) | 2,632人(73.1%) |

これは、いわば「調査の保証書」のようなもので、調査がどのように実施され、結果がどのように入手されたかが端的に分かるような内容になっております。

「1」と「2」については説明がいらぬと思います。「3 配付回収法」というのは「調査員が調査票を、調査相手の方にお渡しし、記入していただいたものを後日、回収に伺う」という方法です。「留置法」と呼ばれることもあります。

「4 調査対象」は母集団を指します。そして「5 調査相手」は、私のこれまでの話ではサンプルと言っていたものです。人間に対してサンプルでは失礼ですので、NHKでは「調査相手」と示します。そして、その人たちは「住民基本台帳」を使用して、「層化無作為2段」と呼ばれる方式でサンプリングしたことを示しています。「層化無作為2段」とは、ランダム・サンプリン

グの1種です。さらに、全国の300の地点からそれぞれ12人をサンプリングして、合計3,600人の方に調査を依頼したことを示しています。最後の「6」は、2,632人から回答を得たこと、そして、2,632人は3,600人に対して73.1%に当たることを示しています。

調査結果には、これらの項目をつけるべきであるというのがNHKの考え方で、それぞれの調査機関によって表記方法は異なりますが、考え方は共通していると思います。

有効率については、のちほど触れますので、ここでは有効数についてお話しします。先ほどの誤差式で分母にあった「n」がこの例では「2,632」です。「50」の2乗が「2,500」ですから、分母は「50」より少し大きな値となります。したがって、「k」以外は「50分の50」、つまり、「1」より少し小さくなりますので「1」と考え、「k」を「2」として誤差は「2」と見積もっておけば安心でしょうか。

こうした計算ができるように、「有効数」は必ず表記すべきです。その他の項目についても表記

しておくべきだということについては、詳しくご説明している時間がありませんので省略しますが、これらの項目が表示されていることによって、調査が誰を対象に、どのように行われたかが明確に分かることはご理解いただけると思います。

10. 調査不能による歪み

次は有効率についてのお話です。例えば、母集団全員を対象に調査を行い、60%の人から回収した場合、この調査結果をどのように利用したら良いでしょうか。有効率が60%しかないのですから、直感的に、調査結果は慎重に利用したほうが良いと思われるでしょう。では、有効率が50%であった場合にはどうでしょう。半分の人からしか回答を得られなかった場合ですね。このような場合、調査結果を利用するのを諦めることができるでしょうか。

有効率、その反対は調査不能率ですが、それらが調査結果にどのように影響するかは、次の計算式で求めることができます。

有効率（調査不能率）の影響

$$p = p1 + (100 - r) \times \frac{p2 - p1}{100}$$

r: 有効率(%)、p: 母集団での比率(%)

p1: 有効者での比率(%)、p2: 調査不能者での比率(%)

この式は、全員から回答を得られなかった場合、どのように計算すれば全員の結果を求められるかを示しています。この式には、先ほどのサンプリング誤差とは異なり、「N」で示した母集団の要素の数や「n」で示したサンプル数が含まれていません。つまり、それらは関係ないということです。したがって、全員を調べた場合にも、一部を取り出して調べたサンプル調査の場合でも適用できます。

では、各変数、「p1」や「p2」について、見て

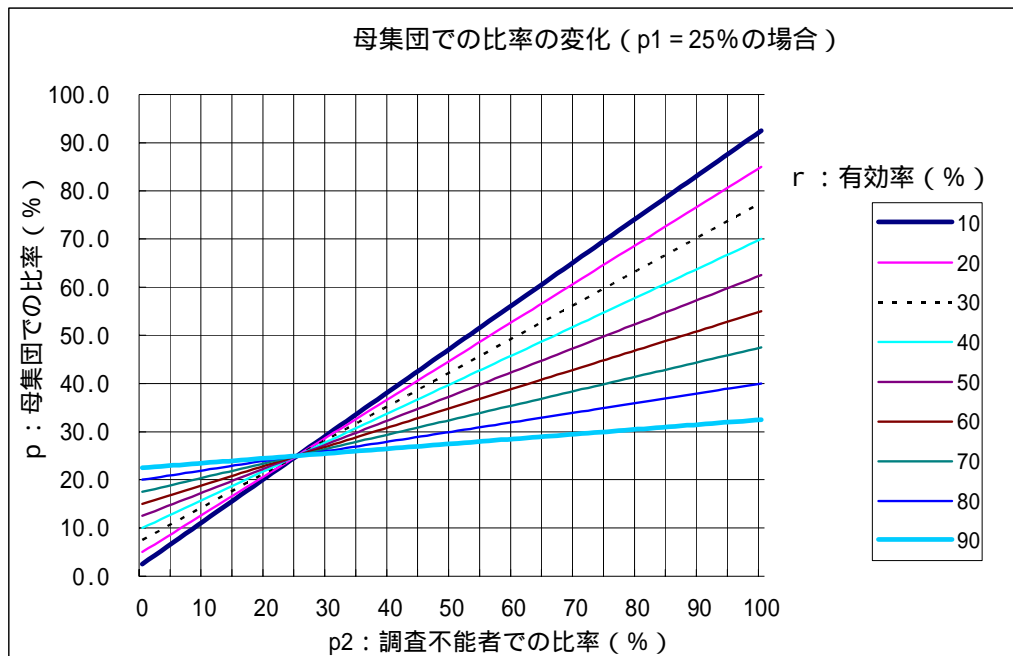
いきます。この中で、「r」つまり有効率と「p1」つまり有効者での比率、割合は調査結果から分かります。ところが、「p2」つまり調査不能者での比率は分かりません。そうすると、母集団での比率を計算する式といっても、結局「p」は計算できないのではないかとと思われるでしょう。

ただし、「p2」つまり調査不能者での比率は、この式では%で示していますが、それは「0%」から「100%」の間にあることは分かっています。調査できなかった人の中で、例えば、ある意見に

賛成する人が一人もいなければ「p2」は「0%」
 ですし、もし全員が賛成していると考えれば
 「100%」です。「p2」は必ず「0%」から

「100%」の間にあります。
 式だけでは分かりにくいので、計算した結果を
 グラフにしました(図3)。

図3 有効率(調査不能率)が調査結果に及ぼす影響



これは、有効者での比率、つまり、調査結果が「25%」であった場合、有効率が変化すれば、母集団での値がどのように変化するかを示しています。線が何本もありますが、それは有効率が10%から90%の場合を想定しているからです。例えば、有効率が90%の場合について見ますと、一番角度が小さい線、あるいは、右端では一番下の線が、有効率が90%の場合です。そして、右端は調査不能者での比率が100%の場合です。縦方向が母集団での比率を示していますから、調査結果、つまり有効者での比率が25%であった時、調査不能者では100%であっても、母集団の比率は30%より少し大きな値を考えれば良いことを示しています。もし、調査不能者では0%であったしたら、左端ですので、20%よりは大きいこととなります。

では、次に有効率が10%の場合について見ていきます。今度は角度が一番急な線、右端では一番上の線です。すると、右端では90%を超えていま

す。有効率が10%しかなかった場合には、調査結果が25%であっても、母集団では、あるいはすべての人から回答が得られた場合に90%以上になっていた可能性があることを示しています。

先ほど、調査結果には「調査の概要」を必ずつけるべきだと申し上げた理由の一つが、これでお分かりになると思います。調査結果を利用するには、サンプリングの方法以外にも、有効率にも注意して利用することが重要です。

11. 新しい調査方法について

最後に新しい技術についてお話ししたいと思います。つまり、電話とインターネットを用いた調査についてです。ただし、新しい技術と申しましても、よく言われるように紙と鉛筆の世界から電子の世界へ変わっただけで、要点は変わりません。つまり、どのようにサンプルを選ぶかという点と、

どのように正確に計測し、記録するかという点です。そして、正確に計測し、記録するという点につきましては本日はお話しする時間がないので、どのようにサンプルを選ぶかという点から電話とインターネットを用いた調査についてお話しします。

まず、電話を用いた調査についてであります。皆さんは「RDD」という言葉を聞いたことがあると思います。「RDD」とは「Random Digit Dialing」の頭文字をとったものであります。このうち「Digit」は「デジタル」と語源が同じであって「数字」と考えれば良いと思います。したがって「RDD」とは「乱数を発生して電話する」というような意味になります。

つまり、先ほどランダム・サンプリングについて説明した中で出てきました「乱数表」を用いた方法と考え方は同じです。現在では、電話を用いた調査はほとんどが「RDD」によるサンプリングで行われている、とされています。

日本でRDD法が普及した背景には、電話帳に番号を掲載していない人が増えたことがあります。電話を持っているのに、電話帳には掲載していない人です。RDD法は、簡単に言えば、使われている可能性のある電話番号を、すべて発生させて、その中からサンプルを選ぶ方式です。しかし、実は、その中には、使われているにも関わらず、含まれない電話番号があります。IP電話と携帯電話です。電話番号で言えばIP電話は「050」で始まるものと、従来の固定電話と同じ番号体系に属するものがあります。このうち「050」で始まるものがRDD法の対象に含めることができない、とされています。また、携帯電話はすべてRDD法の対象外になっているようです。なぜ、対象外なのか、その理由を私は把握しておりませんが、利用者が住んでいる場所の情報が得られないことが大きな影響を及ぼしている

ようです。

さらに、抽出確率の話をしてきましたが、RDD法では対象者が何台の電話を使用できる状態にあるかが、抽出確率に影響を及ぼしますが、それを正確に把握するのは結構大変です。

RDD法は、抽出の過程が何段階にも分かれていますので、それを先ほど紹介したような「調査の概要」の中で端的に示すこととなかなか難しいのですが、調査を実施した側では、できるだけ伝わるように努力すべきです。

一方、調査結果を利用する側では、新しい方法としてRDD法が活用されていますが、課題はまだ多く、また携帯電話もこれからますます普及していくでしょうから、新しいということだけで信用しないで、実際のやり方に注意しながら調査結果を利用することが重要だと思います。

もう一つの新しい技術は、インターネット調査です。

この調査に回答できる人は、パソコンや携帯端末などを所有している人に限られます。これらの機器も急速に普及していますが、それでもお年寄りなどでは、利用していない人が多いことは良くご存知だと思います。

冒頭、ご紹介した「リアルタイム世論調査.net」を世論調査と考えるかどうかは、利用できない人が除外されている点をどう考えるかにもよります。また、インターネット調査の多くは、事前に登録したモニターを対象に行われています。モニターからはランダムにサンプルを選んでいますが、そもそもモニターとはどのような「母集団」であるか、という点も考えなくてはならないことでしょう。

新しい技術としてのRDD法やインターネット調査を、どう評価し、どう活用するか、本日、私のお話が参考になればありがたいと申し上げまして、私の話を終了させていただきます。