

統計手法の利用と誤用

文部省統計数理研究所教授

鈴木 義一郎

三つのウソ

ダレル・ハフという人の書いた『統計でウソをつく法』という本に、統計にまつわる珍談・奇談がいろいろと紹介されている。

……世の中には三つのウソがある。それは口に出して言うウソと、知らん顔をして黙っているウソと、それに統計である。

……統計を使った議論とはその統計を使わなくとも自明のことか、またはどうでもよいことである。

……酔っぱらいは街灯を明かりのためよりも、奇りかかると使っているが、統計も事実を明らかにするためよりも、ごまかすために使われることが多い。

これらの言葉は、いずれも統計の誤用と悪用に對するシンラツきわまりない風刺である。しかし、誤用の裏には正しい使い方が隠されている。悪用の反面には優れた効用も秘められている。要は、それをを用いる人間様次第ということになる。

ところで、この「数」という代物が、いつごろから人間社会に入りこんできたのだろうか。昔、物々交換していた時代、どれとどれとが「等価」であるか、いろいろと比較してみる必要があった。やがて、交換の対象となる品物が増えてくると、個々の品物に「値段」という数値を対応させておくほうが、はるかに便利であることに気がつく。貨幣経済時代の幕明けである。つまり、「数」が賦与されたということは、いろいろなものを比較したいという欲求が先行していた。

どの学校を受験するかを決めるときの指針となる「偏差値」、合格・不合格の決め手となる「総合得点」、就職しようとする会社の「初任給」と、人間万事いろいろな数にとり囲まれて生活している。日本人はまた、算術的能力にたけていると評価されている。ほとんどの日本人が掛け算の「九・九」を空んじていることも、驚異的らしい。

しかしこれも、単に日本語というゴロアワセに適した言葉を使っているからである。むしろ機械的な計算に慣れすぎると、ものごとを深く考えずとも、正しい答えが求められると錯覚しやすくなる。現実の複雑な問題に背を向け、スッキリした数式の展開だけに酔いしれていると、解けるよう設定された問題しか解けないことになる。

ともあれ人間は、いろいろなものを比較する際に、「数」という便利の上もない代物を用いてきた。しかし、この数の背景となっている現象を適正に認識してないと、うっかりして、とんでもない比較をしてしまう危険がある。毒物の許容基準に対し「安全基準」と命名したために、その基準ぎりぎりまでは“安全である”と錯覚してしまったりする。

「比較行為」はなにも、人間だけが行っているものではない。渡り鳥は、より快適な場所を求め移動していくが、寒暖計を眺めて移動日などを決めているわけではない。なまじ人間は、なににつけ「数」に頼って判断するという慣習を身につけてしまったため、数に振りまわされて不適切な比較をしても、気づかずじまいでいることが少なくない。

「比較」は統計の魂である。

外国と比較して、日本の夏は高温多湿で蒸し暑い。ポロシャツに麻のズボンといった軽装に比較して、背広にネクタイというサラリーマン・ルックは、日本の風土に合わない。企業内の一員として、大勢の人と比較して目立ちすぎる服装はマカリナラン。それなら皆がこぞって、もっと涼しい夏服を考案して着用すればいい。

在庫一掃のバーゲン商戦では、「買いどきの今がチャンスと 急がせる バーゲンセールのだましの手口」なのである。夏物をいま手に入れても、次の夏まで生きられるという保証はない。ともあれ比較するときには、比べる対象を明確にしておくことが肝心である。

「ちょっと広めの3LDK」とは、どの位の広さなのか。税率を変えずに控除額だけ引き上げてみても、はたして「減税」と云えるのか。また比較すること自体を罪悪視することから、変な意味での「平等」がまかり通る。教育の機会均等だからといって、勉強する気のまったくない学生が増えていい道理はない。能力差はむしろ、相応に評価されて然るべきと思う。

幼児期に「トンビがタカを生んだ」ともてはやされた子が、成長してタダの人になれば、「カエルの子はカエル」と自嘲する親。また、立派な親の威光で世に出れば「親の七光り」、失敗すれば「不肖の子」。「トナリは何をする人」かが気になり、「トナリの芝生」が青く見え過ぎてイライラ。トナリの息子が一流校へ入ったかどうか気になり、トナリの亭主が課長になったと聞いてはイライラ。「少しぐらい美人だからってナニサ」とつぶやくの

は、決まって美人でない女性。「東大を頂点とする学閥偏重を打破せよ」と息巻くのも、エリートから外れた男性。このように、人間万事、比較せずにはられない。

物価の上昇加減をみるのがお馴染みの消費者物価指数。特定の個人の感覚にだけ合わせては作れないから、ほとんどの人の“実感”からずれることになる。ところで消費者物価指数とは、全ての品目について標準的な消費をする“仮想人間”を対象に算出されたものである。ほとんどの人の支出パターンは、この“仮想人間”のそれとズレている。皆が同じ程度にズレていると感じるから、公平なモノサシといえる。

「同じ服 全ての人に 着せてみよ ピタリ合う人 ひとりもない」のである。物価指数には、教育費だけに関係したもの、ある特定の食品だけのものと、よりきめ細かな数値も算出している。たとえばこの十年間国立大学の授業料は5倍になったが、テレビ等の電気製品は軒並み安くなっている。また地域格差指数でみると、島根では和歌山県の3分の1の値段でまぐろが買える。

美しさでは“姿秀子”だが、気だてとなると“氣立良子”が上だ。お嫁さんとしてどちらを選ぶべきか、迷うところである。複数個の“好ましき”を同時に考えていく場合、各項目ごとに「ウエイト」をつけた足し算を行って、総合的なスコアを算出する必要がある。人それぞれ好みもまちまちだから、このようなスコアのパターンもいろいろと様変わりする。一般に、多元的な情報をひとまとめの数値で表わしたものは「指数」と呼ばれる。嫁さん選びでは、さしずめ“花嫁指数”といったものを定めて比較を行うことになる。

データの氏・素性

結婚相手を吟味するとき、どんな「家」の出かを気にする人が多い。家どうしのつき合いが無視できないお国柄、似た環境に育っていれば、相性もよくなる。世の中には、いろいろな類のデータが出まわっているが、「いつ誰が 何をするため 集めたか 調べて使え 既存のデータ」である。

調査対象者に申告して貰う形式の統計に、ある程度ウソが含まれることは避けられない。昔は、誕生日を1月1日として届ける人が多かったため、1月生まれの比率が異常に高かった。男と女の生まれる比率が、女1に対し男1.06のように、どの年もほぼ安定している。この「出生性比」が、昭和41年に突如乱れた。いわゆる丙午の年で、この年に生まれた女兒を、前後の年にニセ登録したためである。

出産力に関するサンプル調査で、初婚と再婚のパターンで分けした項目をみると、初婚どうしのカップルが94%となっている。ところが『人口動態統計』のほうでみると、初婚どうしの組合せが9割を超える年はまったくない。一般に、被調査者のウソの回答を回避するうまい方法が見当たらない。

大方の人は、金とか権力そして高等理論に弱い。大金をつぎ込んで得られたデータとか、権威筋から発表される報告書、そして超大型計算機で導出された結果と、この類のものほど信用されやすい。だから、「膨大な 金で集めたこのデータ 立派な結論 出てトーゼンよ」といった態度のご仁が多くなる。

“統計数字”が、データ分析の主体であること

は確かであるが、それだけが判断材料のすべてであると思いきまないことである。数字の“背景”となっている現象をよく観察し、その“意味合い”を読みとることが大切である。

情報の一部分を、無意識にせよ意図的にせよ、隠して伝えられるために誤解を生むケースも少なくない。その典型が広告表現によくみられる。○ ○みそは、何とひと味違うのか。一体、何に対してモルト100%なのだろうか。

「デート時は 親切だった好男子 こんなハズ（亭主）では なかったはずよ」と悔やんでみても、あとの祭り。道路交通事故による死者の数が、また1万人を超え続けている。警視庁発表のこの数字、事故発生後24時間以内に死んだ人だけ。その後の死亡者も加算した『人口動態統計』では、1万3千人をはるか超えている。常に、データの氏・素性を確かめておくことが肝要となる。

均してみる

梅雨明け宣言をいつ出すかで、頭を悩ましている気象庁の予報官。「久しぶりに晴れると言ったから、ふとん干してきたのに、どうしてくれんのサ!」。「予報が雨だったから傘もって出たのに、重いったらありゃしない!」。

予報のはずれに対し、多くの人が存外シビアな反応を示す。ところが、晴の予想で結果も晴れというケースもかなりあるのだが、とりたてて問題が起きないことは忘れられる宿命にある。長期間にわたって均してみると予報は結構当たっている、と筆者には思えるのだが……。

夏休みのシーズンになると海へ山へ、そしてふるさとへと、切符の予約をしようとする人々で、

駅はごった返してくる。これでも均してみると利用客は減っているというのが、JR関係者の弁。平年並みの混雑だったら、とても旅行なんかしようとも思わない。

ついに大関にはなれずじまいだった高見山に、大関でもたついている小錦。強いときにはやたら強くても、均してみても強いと判断されない限りは所詮横綱にはなれない。

とかく日本では、個を主張せず集団でものをいうことをよしとするお国柄である。平均を気にし、平均からの離脱を恐れる。海外旅行も集団でしか行けないし、会社が退けて一緒に呑むのも同じ会社の同僚。その後のハシゴ酒まで、場所は違えどメンバー変わらずといった調子である。

しかしこの均質性こそが、経済大国ニッポンを支えてきた源流なのかも知れない。行動パターンが少し違う世代が登場するや、新人類といったレッテルを貼る。不均質性に寛容になることも大事だが、むしろ各人がもっと個を主張するようになったほうが得策ではなかろうか。「平均が 好ましいとは 限らない」といったケースもあるということ、是非認識しておいて頂きたい。

平均体重に満たない幼児を健康診断につれていったら、保健婦さんに「だめじゃないですか、もっとミルクを飲むようにさせなくては」と叱られた。すべての母親が、平均体重以上の子供に育てようとしたらどうなるか。幼児の平均体重はどんどんうなぎのぼりで、まさに天井知らずとなる。「平均的」という形容句を“標準的”と解釈しないで“理想的”と勘違いしてしまっているところに問題がある。平均を上まわる体重の子がいる限り、平均体重以下の子供も同程度にいてちっとも

おかしくはない。

平均は、データの総和をデータ数で割ればよいから、どんなに数学が不得手な人でも、電卓で簡単に求められる。問題は、算出された平均値の“意味合い”をどう解釈するかにある。平均値は、データの分布を質点系の分布のようにみなしたときの“重心”の位置に対応する。

たとえば、分布形が2コブらくだのような形をしているデータの場合を考えてみる。平均値が谷の辺の位置にあるので、データを良く代表する値であるとは認めがたい。山が2か所以上にあらわれるようなケースでは、異種データが混在しているものと考えて、うまく“層別”して説明できないものかと検討を加えてみる必要がある。

とはいえ大抵の場合、平均がデータ全体の様子をよく代表する値であることは、確かである。ただ、データの“中心”に関する情報については遜色がなくても、データの“散布度”に関してまでは言及し得ない。

散布度に関する情報も吸収したければ、標準偏差のようなものを考えてやる必要がある。この値は、各データから平均値までの“ずれ”を差の平方値(平方偏差値と呼ばれる)で評価して平均し、その平方根をとったものである。つまり標準偏差とは、平均値からのずれの“平均”なのである。結局、(算術)平均と標準偏差との2組の数値に縮約する方法が、最もスタンダードなデータの縮約法ということになる。

みそ汁の味見とサンプル調査

筆者のような年配層には、アツカカ〜イみそ汁はどうしても欠かせない。うまいみそ汁をつくる

には、みそを入れたらよくかき混ぜて味見をするのがコツ。この“かき混ぜる”という操作を怠ると、鍋全体のみそ汁が、まるで見当はずれの味になってしまう。

初めての食堂に入って、店頭飾られたサンプルと値段とを見比べる。ところが「サンプルを眺めて決めた トンカツの 出されたものの 小ききことよ」といったケースが多い。ロウ細工師の腕前は、実物より少しだけ豪華にみせるようなものが作れるかどうかで評価される。

「一を調べて十を知る」といったこの種の行為は、就職する会社に関する情報や結婚を前提としたつき合い相手の情報、また本屋でのパラパラメクリといったように、日常茶飯事的なできごとのなかにもたくさんみいだせる。極端に偏った一部分から全体を類推すると、「偏見」ということになる。調べるべき適正な一部分をどう抽出するか、もの見方の根幹に関わる問題である。

総務庁が5年ごとに行っている『国勢調査』。10月1日午前0時における人口数の全貌を把握するための、基本的な「全数調査」である。しかし、一般に行われているのは、「全体の ソックリさんを選びだし 特性調べる サンプル調査」のほうがかほとんどである。調査に要する費用や時間などが、大幅に節約できるからである。よいサンプルとは、いうまでもなく全体とよく似ている一部分のことである。みそ汁の味見の場合のように、よく“かき混ぜた”一部分をとりだすことがポイントである。

そのようなサンプルを選ぶには、全対象集団の数だけ“くじ”を用意し、そのうち調査用の数に相当する“当たり”を作っておく。そして、当た

りを引いた人だけを調査対象のサンプルとみなす。これが、「ランダム・サンプル」と呼ばれているものである。このようなサンプルにもとづく推論が、全体に対する結論から、若干の“ずれ”を生ずるであろうことは、覚悟しておく必要がある。このようなずれのことを、「サンプリングエラー」と呼んでいる。

一般にサンプル数を多くとれば、それだけサンプリングエラーは小さくなる。サンプル調査による誤差がどの程度になるかを見積ることが、実用上大切になる。たとえばサンプル数300くらいから算出されるテレビの視聴率調査、通常3%くらいの推定幅をつけて考えないと正しい値は見積れない。

またサンプリングの設計の段階で振り当てたものがランダムであっても、回収されたサンプルがそうであるとは限らない。サンプルの未回収分によって生じる誤差のことを、「無回答誤差」という。調査を実際に進めていく段階でも、いろいろな種類の誤差が入ってくる。たとえば『出産力調査』における初婚どうしの夫婦の比率が、実際よりも高くなるといったように、聞きとり調査に、ウソ回答をどうしても回避できない。

このような調査対象者に起因する誤差から調査員の犯す誤差、その他調査環境の影響などで生ずる偶発的な誤差まで考えていくと、正に誤差の洪水である。この種の誤差に加え、調査票(特に質問形式)に起因する誤差や集計の際の誤差などを総称して「ノン・サンプリングエラー」と呼んでいる。