

統計調査について

統計調査のむずかしさ

調査によるデータの収集は、統計的処理によるデータ解析の出発点であり、ここが正確でないと後の処理は全く無意味になってしまいます。しかし、不注意な調査は案外いろいろなところにあります。また、わざと不当な調査を行い、都合の良い結論を導く「我田引水」的な調査例も、残念ながらよくあります。

世の中に「アンケート」があふれているように、統計調査というのは簡単にできるように思われているふしがあります。しかし、例えば「質問数が多すぎて、だんだんいい加減に回答するようになってしまう」アンケートには、日常よく出会うものです。

社会調査士資格認定科目には、調査に関する科目が7科目中5科目あり、すべて本学部で開講されています。調査の詳細についてはそちらで学んでいただくとして、今日は「悪い調査」の例をいくつか見てみましょう。

なお、国が行う国勢調査などの統計調査については、調査が「ズル」によって歪められることのないように、統計法という法律によって、きびしい規制・罰則があります。2013年には、愛知県東浦町で、「市に昇格するために、2010年の国勢調査の結果を改ざんして人口を水増しした」という疑いで、当時の副町長が「統計法違反」で逮捕され、有罪となるという事件がありました。また、省庁での統計不正が明るみに出たことがあります。これは「信義を世界に失うがごとき」¹深刻な事態です。

良い質問とは

良い質問とは何かを考えるため、次の「悪い質問」の例を見て、何が悪いのかを考えてみてください。

1. もしあなたが車を運転していて、がまんできないくらい眠くなったら、あなたはどうしますか。

車を運転しない人にとっては、答えようのない質問です。「もし〜だとすれば」という仮定の質問をするには、その仮定に十分な必然性が必要です。

2. あなたは、世の中で一番大切なのはお金だと思いますか。

「あなた自身がお金が一番大切だと思っていますか」なのか、「『世の中の人はお金が一番大切だと思っている』とあなたは思っていますか」なのか、意味があいまいです。

3. あなたは、「学歴よりも実力が大切だ」と思いますか。

「学歴と実力は別のものである」「学歴と実力は比較できる」という前提は、必ずしも正しいとはいえません。

4. あなたは、「宇宙開発は、安全保障上重要なので、進めるべきだ」という意見に賛成ですか。

「宇宙開発は安全保障上重要だが、進めるべきではない」「宇宙開発は安全保障上重要ではないが、それでも進めるべきだ」という意見の持ち主は、この質問に答えることができません。これは、「宇宙開発を進めるべきだ」とする理由は「宇宙開発は安全保障上重要だから」以外にない、と質問者が勝手に思い込んでい

¹1945年8月15日の昭和天皇による「終戦の詔」（いわゆる玉音放送）より「若シ夫レ情ノ激スル所濫ニ事端ヲ滋クシ或ハ同胞排擠互ニ時局ヲ亂リ爲ニ大道ヲ誤リ信義ヲ世界ニ失フカ如キハ朕最モ之ヲ戒ム」

ることが原因です。このように、質問に2つ以上の論点があり、どれを聞かれているのかが不明確な質問を「ダブルバレル質問」といいます²。こういう場合は、まず『宇宙開発を進めるべきだ』という意見に賛成ですか」という質問をした後で、さらにその理由を問うようにします。

5. 当店のお客様に対する対応はいかがでしたでしょうか、次の中からお選びください： (1) 不満
(2) おおむね満足

あきらかに「満足である」という回答に誘導しようとしています。このような選択肢を作りたいのなら、1. 不満 2. やや不満 3. どちらでもない 4. おおむね満足 5. 満足 の5つの選択肢にしなければなりません。

心理的効果を考える

人間の判断は、直観や印象といった心理的効果に大きく左右されます。その影響の例を見てみましょう。

1. 宴会コースのメニューをいくつか設定するとします。店主は、7000円のコースに人気が出てほしいとひそかに考えています。下のどちらにすれば、その意図が実現されやすいでしょうか。

(a) 宴会メニューは、6000円・7000円・8000円・9000円の4通りです。

(b) 宴会メニューは、4000円・5000円・6000円・7000円の4通りです。

最高額や最低額は、選択されにくい傾向があります。予算を7000円にしたいと考えているなら、上の4通りの設定のほうがよいことになります。もちろん、統計調査でこのような細工をするのは好ましくありません。

2. 大学で、学生が講義を評価するアンケートを行います。下の2つの尋ね方で、回答に違いが出るでしょうか。

(a) あなたはこの講義に満足しましたか。次の数字に○をつけてください： 1 (非常に不満)

2 3 4 5 (非常に満足)

(b) あなたはこの講義に満足しましたか。次の数字に○をつけてください： 1 (非常に不満)

2 3 4 (非常に満足)

とくに言いたいことがない場合は、「どちらでもない」に○をつけたいくなるものです。5段階評価の場合は「3」に○をつけることになります。しかし、4段階評価の場合はそれを許さず、満足か不満かどちらかを選べ、と強制していることになります。筆者の大学での経験では、4段階にすると、「どちらでもない」人は「3」に○をつけ、満足な人は「4」、不満な人は「2」、とくに不満な人が「1」をつけるようで、「とくに満足」な人がつける点数がないようです。

誰を調査しているのか

調査をするときは、当然対象となる集団を想定しています。ところが、調査した対象が、想定している集団と違ってしまふことがあります。また、集団の一部を取り出す「標本調査」では「無作為抽出」をする必要があります。無作為抽出については後日の講義で取り上げますが、簡単に言えば「対象となる集団からまんべんなく標本をとる」という意味です。しかし、この場合も集団のうち偏った一部分だけを取り出してしまふことがあります。例を見てみましょう。

² 「ダブルバレル (double-barreled)」とは、元々は2本の銃身をひとまとめでした銃のことをさします。

1. 同窓会の案内を出しました。返事に現在の年収を書いてもらい、同窓生の平均年収を求めました。この結果は信用できるでしょうか？

返事の回収率はおそらく 100%ではなかったでしょう。きちんと案内が届いて返事を返してきた人は、生活が安定している人が多いと思われます。また、収入の少ない人は、見栄を張ってうそを書いているかもしれません。ですから、それを使って求めた平均年収は信用できません。そもそも、こういう個人のプライバシーに関することを、同窓会の案内で尋ねるべきではありません。

2. 日本家屋の伝統的規格では、鴨居の高さは五尺七寸（約 173cm）または五尺八寸（約 176cm）となっています。戦後、日本人の体格が大きく向上したところに、「最近の若い人は背が高い人が増えて、鴨居で頭を打つ人が多い」といわれ、規格を変えなければいけないという声があがりました。

しかし、当時の 20 歳前後の人について、統計データを用いて調べてみると、176cm 以上の人は 5%前後しかいませんでした。なぜ、鴨居で頭を打つ人がたくさんいるような気がしたのでしょうか？

「鴨居で頭を打つ人が多い」という印象を持っている人は、まんべんなく調査した結果がそうなのだという気がしていますが、実際には、鴨居で頭を打たない人は「頭を打って困る」という不平は言いませんから、不平を言う人ばかりが目立って目につくことになります。なお、現在では、鴨居の高さは六尺（約 182cm）以上がふつうだそうです。

3. 第二次世界大戦中、英国軍では、爆撃から帰還したパイロットに調査を行い、「敵攻撃機からの攻撃方向で一番多かったのは後ろ上方からである」という結果が得られました。この結果は信用できるでしょうか？

帰還したパイロットは敵攻撃機からの攻撃を逃れた人であり、攻撃にあつて撃墜された人からは調査していませんから、撃墜された人がたくさんいるのなら、上の結論は信用できません³。

統計資料の読み方

人は、数字で表現された情報はつい盲目的に信用してしまいがちです。しかし、その数字が表している意味をよく知っていないと、適用の限界や結果の意味を誤解し、あやまった結論を信じてしまうことにもなりかねません。統計的手法について勉強する前に、「数字で表された表現」を正しく読む方法を、例を通して見てみましょう。

数字から受ける印象

数字自体には、その値以上の意味はありませんが、人間は自分の常識や感覚と比較することによって、勝手な印象をもつものです。逆に、どのような印象を与えるかを考えて、数字を公表するという手が使われることもあります。

次のような表現について、どのような問題があるかを考えてみましょう。

1. この宝くじには、一等 1 億円が 100 本入っています。

よく宝くじの広告に書いてある表現ですが、そのくじが「何本発売されたか」がわからなければ、100 本の当たりというのが多いのか少ないのかはわかりません。

³なお、この話は広く流布していますが、実際にはもっと詳細な分析が行われています。

2. (阪神高速道路の入り口の表示) 環状線までの所要時間: 70 分

「分」という単位には「短い時間」という印象があります。「70 分」と書いてあるほうが「1 時間 10 分」と書いてあるよりも短く感じます。

3. (a) 2020 年 (令和 2 年) 国勢調査 (総務省統計局) によると, 日本の人口は 126,146,099 人, 65 歳以上の人口は 36,026,632 人である。

(b) 2020 年 (令和 2 年) 国勢調査 (総務省統計局) によると, 日本の人口は約 1 億 2615 万人, 65 歳以上の人口は約 3603 万人である。

よく, 演説で (1) のように細かい数字を並べ立てて, 数字に強いところを見せようとする人がいます。しかし, 1 億人の人口の最後の一桁にどれほどの意味があるのでしょうか? 通常は (2) のような表現で十分です。

何を表現する数字なのか

日常見かける数字は, 現実の何かを数字で表したものです。ですから, その「現実の何か」を誤ると, 全く意味のない数字になってしまいます。

次のような表現について, どのような問題があるかを考えてみましょう。

1. いくつかの食料品販売店を調査したところ, 全体の 3% の商品について, 内容量が表示されている量に足りない, という不正が発見された。したがって, 消費者は平均して 3% 損をしていることになる。

「3%」というのは「商品全体の数のうち, 量目不足があった商品の数」であって, 「量目不足」があった商品が「いくら不足だったか」については何も述べていません。ですから, 消費者がいくら損しているかはわかりません。

2. 1 時間あたりに発生する交通事故件数は, 朝と夕方が多い。朝や夕方の時間帯には, 何か生理的に事故を起こしやすい原因があるのだろうか。

1 時間あたりで考えれば, 朝と夕方は走っている自動車の数が多いわけですから, 事故が多いのは当たり前です。走っている自動車の台数あたりで考える必要があります。

3. 日本での実際のデータによると, たばこを 1 日に 20 本吸う人は, 1 本も吸わない人に比べて 10 倍肺がんで死亡しやすい⁴。したがって, たばこを吸うことは大変危険である。

もしも, 肺がんが非常に珍しい病気で, 1 年間に肺がんで死亡する人が 10 人くらいしかいないのなら, たばこを吸っても吸わなくても肺がんにかかる人はほとんどいないわけで, 「10 倍」といっても大した意味はありません。これが, たばこを吸って肺がんで死亡する人が 100 万人, 吸わなくて肺がんで死亡する人が 10 万人だとしたら, たばこを吸うことは大変危険です。「10 倍」というだけでは, たばこが危険かどうかはわかりません。なお, 実際には, 日本での肺がんによる死亡者は 2021 年の 1 年間で約 7 万 6 千人⁵ですので, たばこを吸うことはかなり危険であるといえます。

統計量を正しく理解しているか

統計量とは, 調査したデータに対して何かの計算をして, データを要約したものです。わたしたちに一番なじみのある統計量といえば, 「平均」でしょう。平均の計算のしかたはよく知っていると思いますが, 「なじみがある」ことを過信すると誤解を生みます。

⁴愛知県がんセンターのウェブサイト「肺がんを防ぐには」より

⁵国立がん研究センターがん対策情報センターの「最新がん統計」より

次のような表現について、どのような問題があるかを考えてみましょう。

1. 気象用語でいう「雲量」とは、快晴を 0、本曇りを 10 として空のうち雲の占める割合を表現する。平均は 5 程度である。したがって、一年間のうち多少雲がある曇りの日が一番多い。

データが「平均値をとる可能性が一番高い」とは限りません。どの値をとる可能性がどのくらいであるかを示す「分布」の違いで、いろいろなパターンがあります。最大値と最小値が一番とる可能性が高く、平均値をとる可能性が一番低い、という分布も考えられます。雲量はこのような分布をとる例です。

2. 「人生 50 年」という言葉があったように、100 年前の日本人の平均寿命は 50 歳程度であった。当時の女性は 5 人 6 人の子供を生むことは当たり前であったので、子育てが終わった後の人生はたった数年しかなかった。

「平均寿命」とは「0 歳児の平均余命」です。昔は乳児死亡率が高かったので、その影響で 0 歳児の平均余命が短くなくなっていました。決して「平均寿命が 50 歳である」からといって「40 歳の人の平均余命が 10 年である」わけではありません。

3. 「500 系のぞみ号」は、運行開始当時、最高速度 300km/h の「世界一速い列車」であった。

「最高速度が世界一速い列車」も「世界一速い列車」の一つですが、実用的にはある地点からある地点へ移動するときの速度が一番速い列車が「世界一速い列車」ではないでしょうか。なお、鉄道技術の世界では、営業列車の速度は「隣接停車駅間の平均速度」で評価するそうで、その意味で「500 系のぞみ」の広島→小倉間は、運転開始当初「世界一速」だったそうです。

今日の演習

1936 年の米国大統領選挙の時の有名な話です。ある雑誌社は世論調査によって民主党／共和党のどちらの候補が当選するかを予想することにしました。そこで、自社の雑誌購読者と電話所有者・自動車所有者の 1000 万人を対象にアンケートを行いました。その結果、共和党の候補が当選すると予想しましたが、見事に外れてしまいました。さて、調査にどのような問題があったのでしょうか？