

2022年度秋学期 統計学 第2回  
統計資料の収集と読み方

浅野 晃  
関西大学総合情報学部



統計調査は、むずかしい🤔

## 今日の話は「悪い調査」

我田引水…  
質問数の多すぎるアンケート…

人はけっして正直ではありません

統計調査の詳細については、  
「社会調査法」等の科目で学んでください。

## 2020年は「国勢調査」の年でした



国勢調査は、  
5年ごとに10月1日に行われます。

10月1日に日本国内に住んでいる人には  
答える義務があります。

答えないと「統計法」違反で  
罰せられます🤔  
(制度上は)

## 統計法違反の罪

### 人口水増し疑い、前副町長を逮捕 愛知・東浦町

2013/2/22付

保存 共有 印刷 印刷 SO ME 印刷 印刷 f その他

市への移行を目指していた愛知県東浦町が、2010年の国勢調査で人口を水増ししたとされる問題で、愛知県警は22日、同町の前副町長、荻須英夫容疑者（63）＝同町緒川＝を統計法違反の疑いで逮捕した。同問題で町側は「事務的ミス」と説明していたが、県警は意図的な水増しだったと判断。組織ぐるみで違法行為が行われたとみて、異例の強制捜査に踏み切った。

日本経済新聞  
2013年2月22日付

### 2013年10月25日 名古屋地裁で懲役4か月・執行猶予2年の判決が確定

近年おきている中央省庁の統計不正は、「信義を世界に失うがごとき」事態。

## 悪い質問の例

## 悪い質問の例

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。

## 悪い質問の例

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。

**「私は車を運転しません」**

車を運転しない人は、この質問に答えられません。

## 悪い質問の例

あなたは、「学歴よりも実力が大切だ」と思いますか。

## 悪い質問の例

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」

…ホントですか？

## 悪い質問の例

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。

## 悪い質問の例

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」

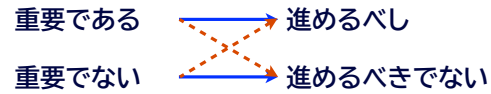
「宇宙開発は安全保障上重要ではないが、それでも進めるべきだ」

…こういう人はどう答えれば？

## 何がいけなかったのか

安全保障上重要？

進めるべきか？



「安全保障上重要なので進めるべきだ」

青い→のパターンしか  
想定していない

赤い破線の→のパターンもある

## こういう質問を

ダブル・バレル(double-barreled)質問という

↑ double-barreled shotgun(銃身が2つある銃)から来ている

2つのもの(「宇宙開発」と「安全保障」)を同時に狙ってはいけない。

心理的効果を考える🧠

## 心理的効果

宴会メニューのコースを設定します。

(🍷7000円のが売れてほしいんだけど…)

(a)6000円・7000円・8000円・9000円 いかが？

(b)4000円・5000円・6000円・7000円 いかが？

## 心理的効果

宴会メニューのコースを設定します。

(○7000円のが売れてほしいんだけど…)

(a)6000円・7000円・8000円・9000円 いかが？

(b)4000円・5000円・6000円・7000円 いかが？

(a)のほうが7000円を選びやすい。人は、端っこの数字は選びにくい。  
(人を特定の回答に誘導する、ズルい質問です)

## 心理的効果

授業評価アンケート

「あなたはこの講義に満足しましたか？」

- 1(非常に不満)・2・3・4・5(非常に満足)

- 1(非常に不満)・2・3・4(非常に満足)

## 心理的効果

授業評価アンケート

「あなたはこの講義に満足しましたか？」

- 1(非常に不満)・2・3・4・5(非常に満足)

- 1(非常に不満)・2・3・4(非常に満足)

下の1・2・3・4のように段階が偶数個だと、「どちらでもない」を選べない

誰を調査しているのか？ 🧑 🧑

## 誰を調査しているのか

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。

## 誰を調査しているのか

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。

返事をくれた人は、そこそこ成功している人でしょう。  
それに、人間は見栄を張るものです。  
そもそも、こういうプライバシーを安易に尋ねてはいけません。

なお、フィンランド<sup>🇫🇮</sup>では、個人の年収は国税庁から公開されます<sup>🤖</sup>  
プライバシーの考えも、文化によっていろいろです。

## 誰を調査しているのか

「最近の若い人は背が高い人が増えて鴨居に頭を打つ人が多い」  
とよくいわれるが、  
全国の統計で調べるとそれほど多くない。なぜ？

## 誰を調査しているのか

「最近の若い人は背が高い人が増えて鴨居に頭を打つ人が多い」  
とよくいわれるが、  
全国の統計で調べるとそれほど多くない。なぜ？

頭をうたない人は文句を言わないから。

## 誰を調査しているのか

爆撃から帰還したパイロットを調査し、  
「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」  
というデータが得られました。  
後ろ上方の防御を固めればいいでしょうか？（第二次大戦中の英国の、有名な実話）

## 誰を調査しているのか

爆撃から帰還したパイロットを調査し、  
「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」  
というデータが得られました。  
後ろ上方の防御を固めればいいでしょうか？（第二次大戦中の英国の、有名な実話）

攻撃されて撃墜された人は帰ってきていませんから、  
調査できていません。  
（※この話は広く流布していますが、本当は、もっと詳細な分析が行われました）

## (参考)「エニグマ暗号機」



フィンランド Museo Militaria(軍事博物館)で浅野が撮影

ひと文字タイプするごとに、歯車が動いて、  
暗号化のキーがどんどん変わっていく  
当時最強の暗号機

その弱点は？

歯車の初期状態を「毎日必ず異なる」設定にしていた  
➡ 一度使ったものを二度と使わないのは、「ランダム」ではない

「柿ピー」の袋の中は、  
柿の種ばかり集まっている場所、ピーナツばかり集まっている場所があって当たり前。

数字から受ける印象 🤔

## 数字から受ける印象

この宝くじには、一等1億円が100本入っています。  
当たりがたくさん入っているのでしょうか？

## 数字から受ける印象

この宝くじには、一等1億円が100本入っています。  
当たりがたくさん入っているのでしょうか？

くじが全部で何本販売されたのかわからないと、  
「100本」が多いのか少ないのかわかりません。

## 数字から受ける印象

- (a)2020年国勢調査によると、  
日本の人口は126,146,099人、  
65歳以上の人口は36,026,632人
- (b)2020年国勢調査によると、  
日本の人口は約1億2615万人、  
65歳以上の人口は約3603万人

どちらの表現が適切でしょうか？

出典：2020年(令和2年)国勢調査結果(総務省統計局)

## 数字から受ける印象

- (a)2020年国勢調査によると、  
日本の人口は126,146,099人、  
65歳以上の人口は36,026,632人
- (b)2020年国勢調査によると、  
日本の人口は約1億2615万人、  
65歳以上の人口は約3603万人

どちらの表現が適切でしょうか？

出典：2020年(令和2年)国勢調査結果(総務省統計局)

日本の人口を説明するのに1の位まで必要でしょうか？

なお、1995年調査では、人口1億2557万人、65歳以上が1826万人  
老人の人口は25年で倍増しました。

2020年国勢調査の  
「人口等基本集計結果」は  
2021年11月30日公表



## 何を表現する数字なのか 12 / 34

## 何を表現する数字なのか

「スーパーでの調査によると、  
全体の3%の商品について、内容量が不足していた。  
消費者は平均して3%損をしていることになる。」

## 何を表現する数字なのか

「スーパーでの調査によると、  
全体の3%の商品について、内容量が不足していた。  
消費者は平均して3%損をしていることになる。」

ここでいっているのは「商品の品目数の3%」であって、  
不足の量がいくらなのかは言っていない

割合を表すときは、  
「分母」が何かをはっきりさせなければいけない

## 何を表現する数字なのか

「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。」

## 何を表現する数字なのか

「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。」

朝夕は、そもそも走っている車が多い。  
「台数あたり」ではどうだろう？

## 何を表現する数字なのか

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。  
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  
たばこを吸うことは大変危険である。

## 何を表現する数字なのか

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。  
たばこを吸うことは大変危険である。

全国で1年間に肺がんで死ぬ人が10人しかいないなら？  
たばこを吸う吸わないにかかわらず、他の病気を心配するほうがよいと思います。

全国で1年間に肺がんで死ぬ人が  
吸う人100万人・吸わない人10万人なら？  
ただちに、たばこを禁止したほうがよいと思います。

日本で1年間に肺がんで死ぬ人は、実際には7.5万人。 国立がん研究センター「最新がん統計」より  
やっぱり、たばこは吸わないほうがよいと思いますよ。

統計量を正しく理解しているか

## 統計量を正しく理解しているか

百年前の平均寿命は50歳程度。  
昔は子だくさんだったので、子育て後の人生はたった数年だった。

## 統計量を正しく理解しているか

百年前の平均寿命は50歳程度。  
昔は子だくさんだったので、子育て後の人生はたった数年だった。

平均寿命とは、「0歳児の平均余命」

子供のうちに死ぬ人が多いと、平均寿命は短くなる。