

2019年度春学期 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。

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 🌐 📄 📧 📧 📧 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

2018年におきた、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、
がまんでこ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。

悪い質問の例

あなたは、「学歴よりも実力が大切だ」
といますか。

「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

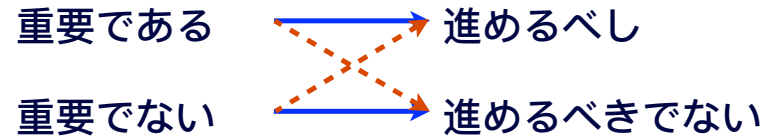
あなたは、
「宇宙開発🚀は、安全保障上重要なので、
進めるべきだ」
という意見に賛成ですか。

「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」

このパターンしか
想定していない

こちらのパターン
もある

こういう質問を

ダブル・バレル(double-barreled)
質問という

(写真略)

http://www.imfdb.org/images/f/fd/Double_barreled_shotgun_1.jpg

↑ double-barreled shotgun

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？

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)

段階が偶数個だと、
「どちらでもない」を選ばない

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。

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、こんなことを安易に尋ねて
はいけません。

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いいません。

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。

攻撃されて撃墜された人は、
帰ってきていません。

(参考) 「エニグマ暗号機」



最強の暗号機の
弱点は…

フィンランド
Museo Militaria (軍事博物館)

歯車を「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

数字にだまされないように🌀

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。

くじは全部で何本販売されたのかがわからないと、「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) 平成27年国勢調査によると、日本の人口は127,094,745人、65歳以上の人口は33,465,441人

(b) 平成27年国勢調査によると、日本の人口は約1億2710万人、65歳以上の人口は約3347万人

出典：平成27年国勢調査結果（総務省統計局）

1の位まで必要でしょうか？

平成7年調査では、人口1億2557万人、65歳以上が1826万人…

何を表現する数字なのか 1234

何を表現する数字なのか

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。

商品の数の3%であって、不足の量がいくらなのかは言っていない

(写真略)
(スーパーの商品棚)

<http://gigazine.net/news/20130323-sugary-soft-drink-death/>

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。たばこを吸うことは大変危険である。

愛知県がんセンター中央病院のウェブサイト「がんの知識」より

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。

1年間に肺がんで死ぬ人が10人なら？

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？

日本で1年間に肺がんで死ぬ人は実際には7万人。国立がん研究センター「最新がん統計」より

統計量を正しく理解しているか

統計量とは

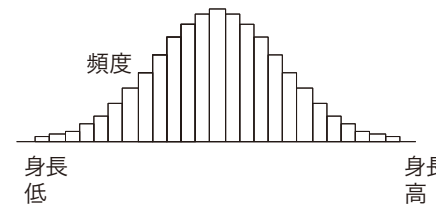
調査したデータに対して
何かの計算をして、
データを要約したもの

一番なじみのある統計量といえば
「平均」 ですが…

統計量を正しく理解しているか

「曇量」は、快晴が0、本曇りが10。
平均は5程度なので、
多少曇がある曇りの日が一番多い。

曇量5の日は一番少ない。



こういうふうには
ばらつくとは
限らない

統計量とは

百年前の平均寿命は50歳程度。
昔は子たくさんだったので、
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。