

2019年度春学期 統計学 第1回
イントロダクション—
統計的なものの見方・考え方について

浅野 晃
関西大学総合情報学部



統計・確率的思考とは何か🤔

数量的思考 📊
微積分的思考 📈
統計・確率的思考 🎲

数量的思考 📊

福島で原発事故

(事故現場写真)

<http://jp.ibtimes.com/articles/17025/20110403/508106.htm>

(沖縄の写真) 沖縄で
480万ベクレルの
放射性ヨウ素！！

<http://okinawa-spot.info/>

記事をよく読むと…

1 平方キロメートルあたり
480万ベクレル

1 平方メートルあたり
4.8ベクレル

人体  にはすでに
4000ベクレル

バナナ  1 本には
20ベクレル

数量的  に考えましょう。

1910年、ハレー彗星接近

(ハレー彗星)

彗星の尾には青酸
が！

http://upload.wikimedia.org/wikipedia/commons/b/b9/Halley%27s_Comet_-_May_29_1910.jpg

(児童書「空気がなくなる日」表紙写真)

<http://www.poplar.co.jp/shop/shosai.php?shosekicode=30200240>

百年前の人を笑えないのでは？

微積分的思考

世の中の数学に関する勘違いは
たいてい



<http://www.benpi-k.com/Baked-banana.html>

(事故現場写真)

原発近くで○ミリシーベルト**毎時**の放射線を
検出

これは**1時間**浴び続けると
レントゲン写真△枚分の被曝に相当...

えっ、レントゲン△枚分の
放射線を浴びたの😱？

原発近くで○ミリシーベルト**毎時**の
放射線を検出

これは**1時間**浴び続けると
レントゲン写真△枚分の被曝に相当...

その量の放射線は一瞬出ただけ
かもしれないし、

そんなところに1時間立っている
わけでもありません。

1 時間あたり100ミリ

の雨 

雨の勢い

○時までの 1 時間に100ミリ

の雨 

実際に降った
雨の量


統計・確率的思考 

放射線障害とは

放射線の粒子が持つ
エネルギーによって
遺伝子に「傷」がつく

(二重らせん
モデル)

http://upload.wikimedia.org/wikipedia/commons/8/81/ADN_animation.gif

傷がつくかどうかは、
偶然  による

(二重らせん
モデル)

傷がついても、
直って病気にならない
こともある

偶然 =

起きるかどうかわからない (人知が及ばない)

確率は

「起きやすさ」を言っているだけ

ある量の放射線を浴びた時、「安全か？」
と言われても困る。

確率を推定する

確率を推定する

ある量の放射線を浴びた時、
病気になる**確率**が
どれだけ大きくなるのか？

ある量以下の放射線は安全、って
おかしいんじゃないの？ 😞

確率を推定する

データを集めて確率を推定するのは、
簡単にいえば

くじびき🎰の結果から
当たり確率を推定すること

そんなこと、できる？

くじのあたり確率

「夏祭り、夜店のくじに当たりなし
露天商の男を逮捕」

(朝日新聞大阪版2013年7月29日)

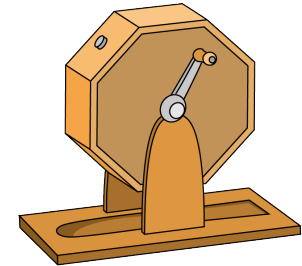
「1万円以上をつぎ込んだ男性が不審
に思い、府警に相談。28日に露店を
家宅捜索し、当たりがないことを確認
した」

半分当たるというくじ

「確率1/2で当たる」というくじを
10回ひいても、1回も当たらなかった

運が悪い🙄のか？

それとも
「確率1/2で当たる」と
いうのがウソ😏か？



こう考える

警察みたいに
全部のくじを調べられないなら、

仮に、本当に「確率1/2で当たる」とする

そのとき、10回ひいて1回も当たらない
確率は、 $(1/2)^{10}=1/1024$

こう考える

本当に「確率1/2で当たる」なら
10回ひいて1回も当たらない確率は
1/1024 (約0.001)

それでも「確率1/2で当たる」を
信じるのは、

確率0.001でしか起きないことが、
いま目の前で起きていると信じるのと同じ

こう考える

確率0.001でしか起きないことが、
いま目の前で起きていると信じる

そりゃちょっと無理がありませんか？ 🤔

というわけで、
「確率1/2で当たる」はウソ、と
考えるほうが自然 「仮説検定」という

確率の計算

ところで、なぜ $(1/2)^{10}$ という
計算になるんですか？

「くじびき」というランダム現象を表す
確率分布モデルを考えているからです

ところで放射線については

ある量以下の放射線は安全、って
おかしいんじゃないの？

病気になった例 (= 当たりくじ) が
数えるほどもないときは、
病気になる確率は「ものすごく小さい」
という不確かなことしかわからない

他の原因でも同じ病気にはなるから、
他の原因と区別がつかない

標本調査と統計的推測

統計的推測とは

「ノルウェー人  の平均身長は、男179cm、女170cmです」

ノルウェー人全員の身長を測ったの？ (リレハンメル五輪のマーク)

<http://www013.upp.so-net.ne.jp/izm-club/norwayhtml>

標本調査

身長は人によって違う (分布している)

ノルウェー人全員ではなく、一部の人だけ (標本) を調べて、分布全体のようなすがわかるか？

くじびきで調べる

一部の人だけ (標本) を調べて、分布全体のようなすがわかるか？

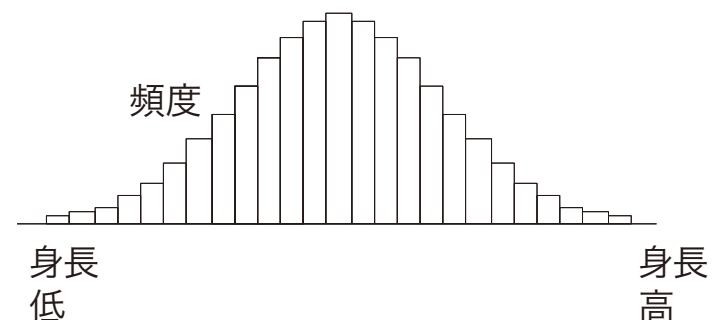
わかります。かなりの程度わかります。

「一部の人」を選ぶのに、くじびきで選ぶ (無作為抽出)

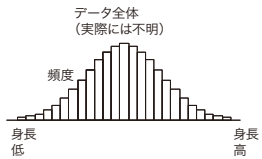
無作為抽出すると

分布がこんなようすのとき

データ全体
(実際には不明)

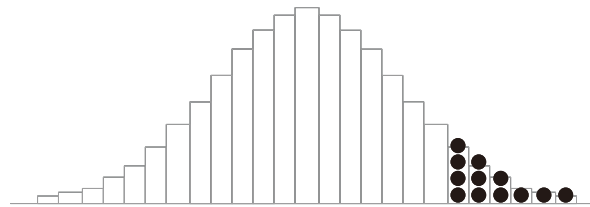


無作為抽出すると



偶然こんな標本 (●) が
選ばれてしまう確率は
小さい

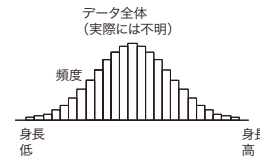
こんな標本が選ばれたら
→大きく偏った推測



2019年度春学期 統計学

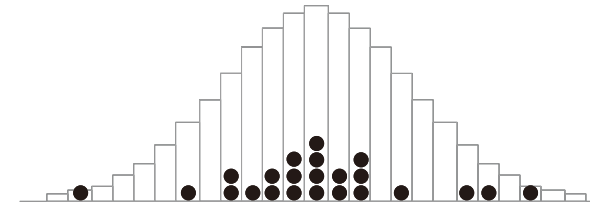
39 - 33

無作為抽出すると



たいていこんなふうに
選ばれる

こんなふうに
標本が選ばれば
→ほぼ間違っていない推測



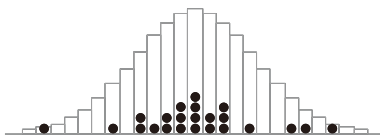
2019年度春学期 統計学

39 - 34

無作為抽出すると

こんなふうに
標本が選ばれば
→ほぼ間違っていない推測

たいていこんなふうに
選ばれる



→標本 (●) を平均すれば
データ全体の平均に
ほぼ近い

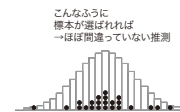
推測できた💡

2019年度春学期 統計学

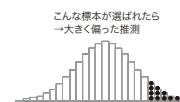
39 - 35

「たいてい、ほぼ」

標本 (●) を平均すれば
データ全体の平均にほぼ近い のはいいとして...



たいていこんなふうに
選ばれるだけであって、



こうなる確率はゼロで
はない (大外し)

いま取り出した標本が
ほぼ当たってるか、大外しかはわからない

2019年度春学期 統計学

39 - 36

区間推定

いま取り出した標本が
ほぼ当たってるか、大外しかはわからない

わからないが、

「ある程度以上はずす確率」は計算できる

「ノルウェー人男性全体の平均身長は、
179cm~182cmの間と推測する。

この推測が当たっている確率は95%」

区間推定

今日のまとめ

確率的思考 🎲

偶然に起きるできごとの
「起きやすさ」を考える

統計学と確率

データをすべて調べきれない場合、

くじびきで選んで調べる

→調べた結果が当たっているかどうか

偶然によるが、当たる確率を述べる

今日の最後に



なぜベンチが
横を向いているの？

転落事故56件を調査

うち33件は
線路に向かって
歩いて落ちていた

思い込みにとらわれず
きちんと調べよう

読売新聞2015.3.31