

1.本当に知りたいのは、どの重さか

導入の狙い

「目的のデータ」ではなく「全体」を知りたい、という動機をはっきりさせる。

ハンバーガーショップのボトニーについて、

- ・今日の30人の重さ
- ・明日の30人の重さ

は、それそれ違います。

では、私たちが本当に知りたいのは、

どの「重さ」でしょうか？

2020/01/アルゴリズム2/09

1/21

1

2

3

2.前回の復習：平均は揃れる

前回の授業では、

- ・標本を取り直すと
 - ・平均が少しずつ変わる
- ことを体験しました。

この「揃れ」は、

間違いではなく、避けられない性質

でした。

2020/01/アルゴリズム2/09

接続の確認

揃れを否定せず、今日の話題の前提として再確認する。

2/21

4

5

6

7

4.標本から1つの数を作る

母集団の平均は、直接調べることができます。

- ・私たちが手にできるのは、
- ・母集団の一部として取り出した
- ・限られた数のデータ（標本）

そこです。

標本を「1つの代表的な数」にまとめる

ことを考えます。

2020/01/アルゴリズム2/09

ここでの位置づけ

この直後に実習②（点推定）における体験を、操作順は大まかに書き下しで示す。

【実習実行・結果作成】

- ・セーターモルヒンに30名を登録
- ・その合計の平均を計算
- ・（計算結果）1つ得られる

4/21

5.その数は毎回同じか？

今作った代表値は、

- ・ある標本（例：30件）から
- ・1回だけ計算した結果

ここで重要なのは、

標本が変われば、代表値も変わらう

という点です。

つまり、代表値は「固定」ではなく、揃れを持つ可能性があります。

2020/01/アルゴリズム2/09

5/21

8

9

10

11

6.なぜ「平均」を使うのか

- ・標本の中には、
 - ・軽いデータ
 - ・重いデータ
- が混ざっています。

平均は、

- ・すべての値を使って
- ・全の重心を表す

代表的な数です。

また、平均には

1つの値に極端に引っ張られにくい

2020/01/アルゴリズム2/09

補足（短く）

「平均」の値は正確で打ち消しあう」という直感を一言で触れる程度に留める。

問題（Q1）：なぜでありますか。

6/21

7.この平均は何をしているのか

ここまで操作を整理します。

私たちは、

- ・標本から代表値（平均）を作り
- ・それを使って
- ・母集団の平均を考えようとしています

このように、

標本から母集団の値を1つの数で推測する

ことを、

点推定

といいます。実習②で計算した平均は、母平均の点推定値です。

2020/01/アルゴリズム2/09

7/21

12

13

14

15

8.点推定の限界

- ・点推定は、
- ・軽く
- ・あまりやさしく

方法です。

しかし、点推定だけでは

どれくらいずらせるか（誤差の大きさ）が分かりません

という限界があります。

次は、

ズレの大きさも一緒に示す方法

区間推定

2020/01/アルゴリズム2/09

次の実習③への橋

この直後に実習③（区間推定）を入れる。

【実習実行・結果作成】

- ・実習②の標本を残す・標準偏差を計算
- ・母平均の区間推定を計算
- ・標準偏差の区間推定を計算

8/21

9.点推定だけでは足りない

点推定では、標本から1つの数を作りました。

しかし、その数には次の弱点があります。

- ・標本が変わればも変わる（揃れる）
- ・それでも「どれくらい揃れるか」が分からぬ

つまり、

推定値の「不確かさ」を表せていない

のです。

2020/01/アルゴリズム2/09

位置づけ

点推定の限界を「ズレの大きさが不確」として再確認するスタイル。ここではまだ「区間推定」などの用語を出さず、「不確かさを表したい」という目的だけを明確にする。

16

17

18

19

10.不確かさは「幅」で表す

狙い

「区間推定=範囲で言う」という直感を内に。ここではまだ計算方法を出さない。

ある幅（範囲）で示す

11.区間推定とは何か

標本から母集団の値を「範囲」で推測する。

それに対して区間推定は、

母集団の値を「範囲」で推測する

方法です。

この範囲を

推定区間（信頼区間）

と呼びます。

用語導入

ここで初めて「区間推定」「推定区間（信頼区間）」の用語を出す。「点→区間」の対比を用いて定義する。

2020/01/アルゴリズム2/09

10/21

2020/01/アルゴリズム2/09

11/21

20

21

22

23

12.なぜ「範囲」が作れるのか（根拠）

区間推定ができるのは、標本平均の扱い方に規則があるからです。

- ・標本平均は真人平均に集まりやすい
- ・極端に外れた平均は出にくい

この規則性（第7回で体験した性質）が、

「どの範囲なら普通か」を決める土台

になります。

902020. アルゴリズム 2.09

12 / 21

横渡し

CFILINの詳しい説明は別スライド（後のグループ）で回収する範囲。ここでは「規則があるから範囲が作れる」までを短く言います。

13.区間推定で言えること

区間推定をすると、推定結果は

1つの数+その周りの幅

になります。

- これにより、
- ・点推定だけでは分からなかった「不確かさ」
- ・推定値がどれくらい信頼できそうか

を、文字で説明できるようになります。

組い

区間推定を「計算の手順」ではなく「言える内容が理屈」のものとして位置づける。この他で計算式へ入る。

24

25

26

27

14.区間推定で「標準偏差」が必要な理由

区間推定では、点推定（標本平均）に幅を付けて「ありそうな範囲」を作ります。

この幅を決める材料は、主に2つです。

- ・ばらつきの大きさ（標本標準偏差 s ）
- ・標本の大きさ（件数 n ）

標準偏差が大きいほど、平均は揺れやすいので

区間は広くなる

逆に標準偏差が小さいほど

区間は狭くなる

のです。

902020. アルゴリズム 2.09

16 / 21

組い

区間推定の「幅」は偶然ではなく、データのばらつき（標準偏差）と件数で決まると言証する。式は次スタイルで最も簡単だけ出る（または口頭で示す）。

15.幅は「標準誤差」で決まる

標本平均の誤れやすさは、次の量で表されます。

$$\text{標準誤差 } SE = \frac{s}{\sqrt{n}}$$

- ・ s が大きい（ばらつき大） $\rightarrow SE$ も大 \rightarrow 区間が広い
- ・ n が大きい（件数多） $\rightarrow \sqrt{n}$ が大 $\rightarrow SE$ 小 \rightarrow 区間が狭い

区間推定は、ざっくり言うと

$$\text{平均} \pm (\text{何個か分の } SE)$$

という形で作られます。

補足（式の扱い）

ここでは「覚える公式」にしない、幅=ばらつき×標本の数×（平方根）という意味だけを理解させる。95%などの信頼（信頼）は、次の実習でツールに任せてよい。

28

29

30

31

16.実習②：区間推定を計算する

ここからは、区間推定を実際に計算します。

今日ここで行なうことは、次の2つです。

- ・標本平均を確定する（点推定）
- ・その平均の「ズレの大きさ」を考へ、範囲を作る

母平均を「1つの数」ではなく「幅をもった範囲」で表す

902020. アルゴリズム 2.09

16 / 21

実習②（区間推定）の仕様

【目的】点推定に「標準偏差」を組み、母平均を範囲として表現できるようにする。

【使用データ】実験3で販売したトマトの平均の標本標準偏差を計算。母平均を計算するためのデータは、前回の実験（トマトの平均は「標準偏差」）によって決まる。同じn=3でも、ばらつきが大きい時はばらつきが大きい

【手順】標準偏差を計算→別のトマト30個を取る→平均・標準偏差・区間の変化を比較する

【時間目安】約1分

17.区間推定の読み方

推定区間が

148 g ~ 152 g

のように読められます。

このとき、私たちが言いたいのは

母平均はこの範囲にあります

ということです。

点推定（150g）だけのときよりも、どれくらいズレそうかが一緒に分かるようになつています。

口頭補足（注意点）

「必ずこの中にいる」とは言わない。母平均（98%など）の範囲の意味は次の幾つか回してよい。ここでは「箱をもって丁寧する」この範囲を強調

32

33

34

35

18.点推定と区間推定の整理

今日ここまでで、母平均の推測には2つの形があると分かりました。

- ・点推定：母平均を1つの数です
- ・区間推定：母平均を幅のある範囲で表す

区間推定によって、

推測の「不確かさ」も含めて判断できる

ようになります。

902020. アルゴリズム 2.09

16 / 21

次回への接続

次回は、この「ズレの大きさ」を基準にして「どれくらい外れたら外れているのか？」を判断する区間推定に進む。

19.ズレの大きさに「基準」を与える

区間推定では、

どれくらいのズレまでを「普通」と考えるか

を、あらかじめ決めます。

よく使われる基準の1つが、95%です。

これは、

- ・標本のばら方をえたとき
- ・その範囲に入ることが多い

という意味で使われます。

口頭補足（信頼水準）

・95%は「標本」ではなく、標本の「基準の範囲」・95%に外れる範囲は広くなりすぎる・90%や99%など、目的に応じて選ばれる。ここでは「ズレの基準を教えて貰っている」ことを理解させる

36

37

38

39

20.この回の到達点

今日の授業で、次のことができるようになりました。

- ・標本から母平均を点で推測できる
- ・ズレの大きさを考へ、区間で推測できる
- ・そのズレに基準（95%）を与える

これにより、

「どれくらい外れたら外れているか」を考へる準備

が整いました。

902020. アルゴリズム 2.09

16 / 21

次回への接続

次回は、この区間やズレの大きさを相棒にして、「この結果は異常か？ それとも違うか？」を判断する区間推定に進む。

21.まとめ：なぜ推測できるのか

母集団は、直接見ることができません。

それでもたたは、

- ・標本を取る
- ・そのばら方を理解し
- ・数で基準を決める

ことで、母集団について考へてきました。

推測統計とは、標本の「報れ」を前提にして判断する方法

授業の閉じ（口頭）

・「当てる」ではなく「範囲をもって判断する」、この考え方方が次の区間推定につながる・今日は計算よりも「考え方」が主役だったことを強調

40

41

42

43