

09 推測統計の考え方（標本と母集団）

1. 本当に知りたいのは、どの重さか

ハンバーガーショップのポテトについて、

- ・ 今日の 30 人分の重さ
- ・ 明日の 30 人分の重さ

は、それぞれ違います。

では、私たちが本当に知りたいのは、

どの「重さ」でしょうか？

導入の狙い

「1回のデータ」ではなく「全体」を知りたい、という動機をはっきりさせる。

2. 前回の復習：平均は揺れる

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「揺れ」は、

間違いではなく、避けられない性質

でした。

接続の確認

揺れを否定せず、今日の話題の前提として再確認する。

3. 今日のゴール

今日の授業では、

- ・ 標本から母集団を考えるとどういうことか
- ・ 1つの数（点）で推測する方法
- ・ 幅をもたせて推測する方法

を学びます。

「ズレ」を前提にして判断する

ことが、今日のテーマです。

導入の締め

「計算が目的ではない」ことをここで一度強調。

4. 標本から1つの数を作る

母集団の平均は、直接調べることはできません。

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）

そこでまず、

標本を「1つの代表的な数」にまとめる

ことを考えます。

ここでの位置づけ

この直後に実習①（点推定に向けた体験）を入れる。

操作手順はスライドに書かず口頭で指示する。

【Excel 実習・簡単仕様】

- ・全データからランダムに 30 件を抽出
- ・その 30 件の平均を計算
- ・「代表値が 1 つ得られる」ことを体験させる

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、代表値も変わりうる

という点です。

つまり、代表値は「固定」ではなく、**揺れを持つ**可能性があります。

実習①の回収

実習①で得た平均値を例に、

「別のランダム 30 件に変えると値が変わる」ことを短く口頭確認する。

ここでは点推定という用語はまだ出さない。

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
- ・ 全体の中心を表す

代表的な数です。

また、平均には

1つの値に極端に引っぱられにくい

補足（短く）

「平均との差が正負で打ち消し合う」という直感を一言で触れる程度に留める。

理論（CLT）は後で回収する。

7. この平均は何をしている数か

ここまでの操作を整理します。

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①で計算した平均は、**母平均の点推定値**です。

ここで初めて用語を確定

「さっきの平均は点推定だった」と後づけで意味づけするのがポイント。

先に体験→あとで命名（第7回の作り方と同じ）。

8. 点推定の限界

点推定は、

- ・ 手軽で
- ・ わかりやすい

方法です。

しかし、点推定だけでは

どれくらいズレるか（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

次の実習②への橋

この直後に実習②（区間推定）を入れる。

【Excel 実習・簡単仕様】

- ・実習①の標本を使い標準偏差を計算
- ・母平均の信頼区間を算出

操作手順は口頭で指示する。

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。

- ・ 標本が変われば値も変わる（揺れる）
- ・ それでも「どれくらい揺れるか」が分からない

つまり、

推定値の「不確かさ」を表せていない

のです。

位置づけ

点推定の限界を「ズレの大きさが不明」として再確認するスライド。ここではまだ「信頼区間」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。

10. 不確かさは「幅」で表す

推定値が揺れるなら、1点だけで言い切るのではなく、

ある幅（範囲）で示す

ほうが自然です。

例えば、次のような言い方です。

- ・ 「母平均は 150g くらい」ではなく
- ・ 「母平均は 148g~152g のあいだにありそう」

狙い

「区間推定＝範囲で言う」という直感を先に作る。ここではまだ計算方法を出さない。

11. 区間推定とは何か

標本から母集団の値を **1つの数** で推測するのが点推定でした。

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」の対比を短い文章で固定する。

12. なぜ「範囲」が作れるのか（根拠）

区間推定ができるのは、**標本平均の揺れ方に規則がある**からです。

- ・ 標本平均は真ん中付近に集まりやすい
- ・ 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。

橋渡し

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

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用できそうか

を、文章で説明できるようになります。

狙い

区間推定を「計算の手順」ではなく「言える内容が増える」ものとして位置づける。この後すぐ実習②で計算へ入る。

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。

狙い

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）と件数で決まることを言語化する。式は次スライドで最小限だけ出す（または口頭で示す）。

15. 幅は「標準誤差」で決まる

標本平均の揺れやすさは、次の量で表せます。

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

- ・ 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÷件数の効果（平方根）という意味だけを理解させる。95%などの係数（t 値）は、次の実習でツールに任せてもよい。

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。

今日ここで行うことは、次の2つです。

- ・ 標本平均を確認する（点推定）
- ・ その平均の「ズレの大きさ」を考え、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

実習②（区間推定）の仕様

【目的】点推定に「不確かさ」を加え、母平均を範囲として表現できるようにする。

【使用データ】・実習①で固定したランダム 30 件の標本

【Excel 操作（詳細は口頭）】・標本平均 \bar{x} を計算・標本標準偏差 s を計算・件数 $n = 30$ を確認・これらを用いて母平均の信頼区間を算出

【重要な観察ポイント】・区間の幅は「標準偏差」によって決まる・同じ $n=30$ でも、ばらつきが大きい標本ほど区間は広くなる

【ミニ確認（時間が許せば）】・別のランダム 30 件を取り直し、平均・標準偏差・区間幅の変化を比較する

【時間目安】約 15 分

17. 区間推定の読み方

推定区間が

148 g ~ 152 g

のように得られたとします。

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そうだ

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っていきます。

口頭補足（注意点）

- ・「必ずこの中に入る」とは言わない・信頼水準（95%など）の厳密な意味は次の塊や次回で回収してよい・ここでは「幅をもって主張する」ことの価値を強調

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つの数で表す
- ・ **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。

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？」を判断する仮説検定に進む。

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、

- ・ 標本の揺れ方を考えたとき
- ・ **その範囲に入ることが多い**

102020という意味で使われます。

口頭補足（信頼水準）

・95%は「絶対」ではなく、判断のための約束・100%にすると区間は広くなりすぎる・90%や99%など、目的に応じて選ばれる・ここでは「ズレの基準を数で決めている」ことを理解させる

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
- ・ ズレの大きさを考え、**区間** で推測できる
- ・ そのズレに **基準 (95%)** を与えられる

これにより、

「どれくらい外れたら外れているか」を考える準備

が整いました。

次回への接続

次回は、この区間やズレの大きさを根拠にして、「この結果は偶然か？ それとも違いがあるのか？」を判断する仮説検定に進む。

21. まとめ：なぜ推測できるのか

母集団は、直接見ることはできません。

それでも私たちは、

- ・ 標本を取り
- ・ その揺れ方を理解し
- ・ 数で基準を決める

ことで、母集団について考えてきました。

推測統計とは、標本の「揺れ」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「根拠をもって判断する」・この考え方が次回の仮説検定につながる・今日は計算よりも「考え方」が主役だったことを強調