

02 統計で何ができるか

1. 今日のゴール：統計を「使う」ための第一歩

今日から、統計の学習を始めます。

統計とは、身の回りで起きていることを、数字を使って整理し、考えるための道具です。この授業では、いきなり難しい計算を学ぶのではなく、「統計を使うとはどういうことか」を体験しながら理解していきます。

本日の到達目標

1. 統計の役割：なぜ「数字」で説明するのかを知る
2. データの形：コンピュータが扱いやすい「正しい表」を理解する
3. 第一步：Excel でデータの全体像をつかむ

統計は、計算ではなく「判断を助ける道具」です。

導入

「今日は統計の最初の日です」と明確に宣言する。数学が苦手でも大丈夫だと安心させてから目標を示す。

2. 統計学の役割：バラバラな記録を「情報」に変える

統計学とは、大量にある「事実の記録（データ）」を、意味のある「情報」に変換する技術です。

- ・要約：たくさんのデータを、平均値などの「1つの数字」で表現する。
- ・判断：新しい施策に効果があったのか、偶然なのかを見極める。
- ・予測：過去の傾向から、将来のリスクや可能性を予見する。

経験や勘（「たぶんこうだろう」）を、客観的な根拠（「データがこう示している」）に置き換えることが統計学の目的です。

統計の役割

「勘 vs 根拠」の対比を明確にし、ビジネスや研究での必要性を伝えます。

3. データとは「事実の記録」である

「データ」とは、目の前で起きた出来事を記録したものです。数値だけでなく、言葉や記号も大切なデータになります。

データの具体例

- ・ **数値**：売上金額、気温、テストの点数、ポテトの「重さ」
- ・ **文字・記号**：商品名、性別、合否、味の「印象」（薄い/濃い）

数値だけでなく、アンケートの回答や「はい/いいえ」といった記号も、集めて整理すれば立派な分析対象になります。

データの定義

「数字以外もデータである」ことを認識させ、分析の幅を広げます。

4. データの種類：質的データと量的データ

データは大きく 2 種類 に分かれます。「計算してよいか？」がポイントです。

質的データ（カテゴリ）

- ・ 分類・区別のためのデータ
- ・ 例：曜日、味の印象
- ・ 基本は数える（件数・割合）

量的データ（数値）

- ・ 大きさ・量を表すデータ
- ・ 例：重さ_g、身長、壳上
- ・ 計算して要約（平均・最大など）

同じ「列」でも、何をしたいかで扱いが変わる。

尺度水準

難しい用語は出さず、「計算して意味があるか？」で判断させる。

5. 検証の舞台：駅前のハンバーガーショップ

「私」は駅前のハンバーガーショップをよく利用しています。特にお気に入りはポテト M サイズです。

- ・ 公表値：ポテト M サイズは 135g とされている
- ・ 現場：すべてを毎回計量するのは難しく、目分量になりやすい
- ・ 疑問：本当に 135g と言えるのだろうか？



舞台設定

「公表値 135g」だけを確定情報として提示する。 \pm の許容はこの時点では出さず、後の議論（検定）に回す。

6.【事件発生】SNS の噂と自分の直感

ある日、「私」はSNSで気になる投稿を見かけました。



- ① 駅前店のポテト、135g より少ない気がする」
- ②「味が日によってバラバラ」
- ③「店員によって重さが違う、味も...」

- ・ **自分の実感**：「確かに、昨日買ったときもスカスカだったような…」

「本当のところはどうなんだろう？」という素朴な疑問が、すべての始まりでした。

導入のポイント

学生に「これ、皆さんも経験ありませんか？」と問いかけ、当事者意識を持たせてください。ネットの書き込みをただ眺める客から、真実を確かめる調査員への視点切り替えを行います。

7.【行動】1週間の徹底調査（S00 の記録）

「私」はうわさを確かめるため、実際に1週間毎日ポテトを購入して測ってみることにしました。



【調査ルール】

- ① 毎日 M サイズを 1 袋買う
- ② キッチンスケールで重さを測る
- ③ 味や調理状態も記録する

・集まった証拠：月曜日から日曜日まで、合計 7 袋分の生データ。

データの信頼性

「感情的に文句を言うのではなく、まず記録（データ）を取る」という、統計的思考の第一歩を強調します。この7日分のデータが、後の実習で使用する「S00」のデータそのものであることを伝えます。

8.【結果】2gの差は判断できるのか？

1週間分のデータを集計すると、次のようになりました。

| 曜日 | 月 | 火 | 水 | 木 | 金 | 土 | 日 | 平均 |
|----|------|------|------|------|------|------|------|------|
| 重さ | 130g | 138g | 125g | 140g | 132g | 129g | 137g | 133g |

- ・ 135g を超える日もあるが、平均は 133g
- ・ 公式値より 2g 少ない

しかし――

- ・ この 2g は問題と言えるのか？
- ・ それとも、たまたま起きた誤差なのか？

この時点では、まだ判断できない

第9回への種まき

この「 $2g$ の差をどう解釈するか」がこの授業のメインテーマであることを告げます。この時点では「 $2g = \text{誤差}$ 」と思う学生と「 $2g = \text{事件}$ 」と思う学生に分かれるはずです。その直感を、12回かけて「論理」に変えていくことを約束してください。

9. ここからは「『私』の調査結果を検証する」

これまで SNS の評判や個人の感想といった「主観的な情報」を扱ってきました。ここからは、「私」が実際に店舗へ足を運んで集めた「データ」を使って、客観的に事実を確かめていきます。

- ・ **統計の役割**：「私」の感覚を信じるのではなく、目の前の「数字」を使って事実を確かめるための道具です。
- ・ **実習の目的**：「私」が記録したデータから、Excel を使ってお店の実態を論理的に導き出す手法を学びます。

「統計は、『私』の調査結果を数字で裏付けるための道具」という意識を持って取り組みましょう。

実習への切り替え

ここから PC 操作が始まります。

自分のデータファイル（poteto1.xlsx）を開くよう指示し、授業のモードを切り替えてください。

10. 今日使うデータ：「私」が調査した「1週間の記録」

実習では、「私」がこの1週間、店舗で計測して入力した Excel ファイル (poteto1.xlsx) を使用します。

- ・ 1行 = 1日分：「私」が計測した7日間の生データです。
- ・ 項目（列）の確認：
 - ・ 曜日：計測した日（月～日）。
 - ・ 重さ_g：自分で測ったポテトの重量。※数値データ
 - ・ 味の印象・メモ：その時の感想。※文字データ

本日のポイント：今日は、計算が可能な「重さ_g」の数値データのみを扱います。味の印象やメモは、今は使いません。

データの前提

学生にファイルを開かせ、特に「重さ_g」の列が計算対象であることを認識させてください。

11. 実習(1)：計算する前にデータを「眺める」

Excelで関数を入力する前に、まずは「私」のデータをじっくり眺めてみましょう。数字を直接見ることで、多くの気づきが得られます。

- ・ 「私」のデータを探してみよう

1. 7日間の中で、一番「重かった日」と「軽かった日」はいつですか？
2. 目標の 135 g を超えている日は、「私」のデータの中に何日ありましたか？
3. 「私」の7日間の結果だけを見て、お店に対して何を感じますか？

計算機（Excel）に頼り切る前に、「私」の目で「アタリ」や「ハズレ」を実感することが大切です。

直感の言語化

いきなり関数を打たせず、まずは自分のデータを眺めて「木曜日は多いな」「水曜日は少ないな」といった特徴を探す時間を設けてください。

12. 平均とは「全体の平らな姿」を見ること

「平均」は、バラバラな重さのデータを「もし全部同じだったら？」と平らにならしたときの値です。

- ・ 数式の意味：

$$\text{平均} = \frac{\sum_{i=1}^n x_i}{n}$$

※ 「『私』の7日間の合計重量を、測った日数 ($n = 7$) で割る」ことで、1日あたりの標準的な量を算出します。

- ・ 平均で見えるもの：「私」の調査における「全体的な重さの目安」。
- ・ 見えなくなるもの：「すごく重かった日」や「軽すぎた日」の個別の体験。

平均は便利ですが、これだけでは「日ごとの差（ばらつき）」は見えなくなります。

意味理解

「足して割る」という操作だけでなく、「平らにならす」というイメージを強調してください。

13. 実習(2) : Excel で「平均」を計算する

「私」のデータの「平均重量」を、Excel の関数を使って計算しましょう。

- ・ 使用する関数 : AVERAGE (アベレージ)

- ・ 操作手順 :

1. 答えを表示したいセルを選択する。
2. =AVERAGE(と入力する。
3. 「私」のデータの「重さ_g (7日分)」をマウスで選択する。
4. Enter キーを押す。

今回の「私」の結果 : 計算すると **133.0 g** になりましたか？ 目標の 135 g に近いでしょうか。

操作成功体験

poteto1.xlsx のデータ通りであれば 133 になります。学生が正しく範囲選択できているか確認してください。

14. 最大・最小とは「体験の幅」を見ること

平均（中心）だけでなく、「データの端（はし）」を見ることも非常に重要です。

- ・ **最大値 (MAX)**：「私」がこの1週間で体験した「最高のアタリ」。
- ・ **最小値 (MIN)**：「私」がこの1週間で体験した「最悪のハズレ」。

キーメッセージ：お客様は「平均」を食べているのではありません。その日の「1袋」がすべてです。平均が 133 g あっても、最小値が 125 g ならば、「私」は不満を感じるかもしれません。

ビジネス視点

「平均さえ良ければいい」という考え方の落とし穴を説明します。

15. 実習(3)：Excel で「最大・最小」を計算する

平均の時と同じ手順で、「私」のデータの「端っこ」を数値化しましょう。

- ・ 使用する関数：
 - ・ 最大値：MAX（マックス）
 - ・ 最小値：MIN（ミニマム）
- ・ 手順：平均の時と同じ「7日分の範囲」を選択して計算します。

今回の「私」の結果：最大は **140 g**、最小は **125 g** になりましたか？

操作の定着

関数名が変わるだけで手順は共通であることを伝え、操作に慣れさせます。

16. 中央値とは「並べた時の真ん中」を見ること

平均とは別の角度から「真ん中」を探すのが、中央値（メディアン）です。

- ・ 定義：データを小さい順に並べたとき、ちょうど真ん中にくる値です。
- ・ 位置の計算：

$$\text{位置} = \frac{n+1}{2} = \frac{7+1}{2} = 4\text{番目}$$

※「私」のデータを軽い順に並べて、4番目にくる重さが中央値です。

- ・ メリット：1日だけ「山盛りのサービス（外れ値）」あっても、その影響を受けにくい指標です。

「平均と中央値は役割が違う」ことを意識しましょう。

概念の対比

平均は合計に引きずられ、中央値は順番を重視するという違いを強調します。

17. 実習(4)：Excel で「中央値」を計算する

Excel では、自分で並べ替え作業をする必要はありません。

- ・ **使用する関数**：MEDIAN（メディアン）
- ・ **操作**：範囲を選択するだけで、Excel が裏側で自動的に並べ替えを行い、真ん中の値を抽出します。

今回の「私」の結果：中央値は **132.0 g** になりましたか？ 平均値 (133.0) と少しズレがあることがわかります。

操作の拡張

MEDIAN の綴りを間違えないよう、スライドを示して確認させてください。

18.「私」の計算結果を並べて比較してみよう

Excelで算出した4つの指標（代表値）を一つの表にまとめます。

- ・ 比較する「私」の数字：

- ・ 平均：133.0 g
- ・ 中央値：132.0 g
- ・ 最大：140 g／最小：125 g

- ・ 問い：

- ・ どれが「一番普通（いつもの量）」を表していると感じますか？
- ・ 目標の135 gと比べて、「私」の調査結果はどう評価できますか？

自分の計算結果を多角的に見ることで、お店に対する「『私』の主張」の根拠が生まれます。

整理

自分のデータを客観的な数値で整理させ、「言えそうなこと」を再考させます。

19. この時点で、何が分からぬ?

計算は正しくできました。しかし、この結果だけで「この店のポテトは足りない」と断定して良いでしょうか。

- ・ 「私」のデータの限界：

- ・ データは「『私』一人」が測った、7日分だけです。
- ・ たまたま「私」が買った1週間が特殊（ハズレが多い週）だった可能性はありませんか？

結論：数字は出ましたが、まだ「お店全体を判断」するには証拠が足りません。客観的な結論を出すには、もっと多くの調査結果が必要です。

到達点の明確化

「自分のデータ（n=7）だけでは判断できない」という気づきを促します。

20. では、どうすれば判断できる？

「私」の7個のデータでは偶然かもしれません。判断を確かなものにするには、どうすればいいでしょうか。

- ・ **解決策**：データの「数（サンプルサイズ）」を増やす。
- ・ **次回の予定**：クラス全員のデータを合体させてみよう。

「私」のデータは小さくても、クラス全員分集めれば、100個以上の巨大なデータになります。

「みんなのデータを合わせれば、真実が見えてくる」

次回の導線

各自の調査結果が、次回の全体分析に繋がることを示唆して締めくくります。