



# 誤差と不確かさ

## ～標準偏差～

標準偏差はExcelで、ちゃちゃっと導出できますが、  
**標準偏差の意義を学んでおくことは重要です。**



speaker

〇〇という分析結果が得られました。  
この数値から××であると考えられます。



audience

その値の**誤差**はどのくらいですか？

この質問における誤差とは



speaker

〇〇という分析結果が得られました。  
この数値から××であると考えられます。



audience

その値の**不確かさ**はどのくらいですか？

**不確かさであり**



speaker

〇〇という分析結果が得られました。  
この数値から××であると考えられます。

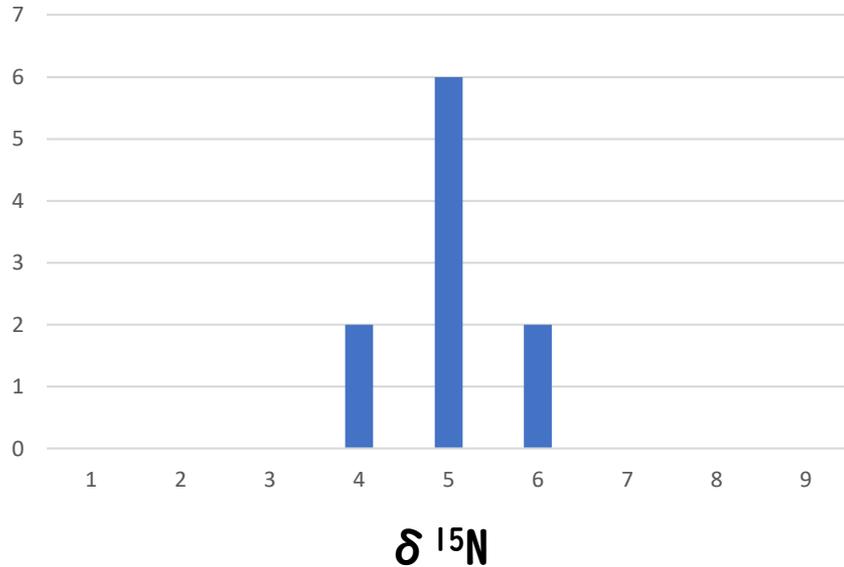


audience

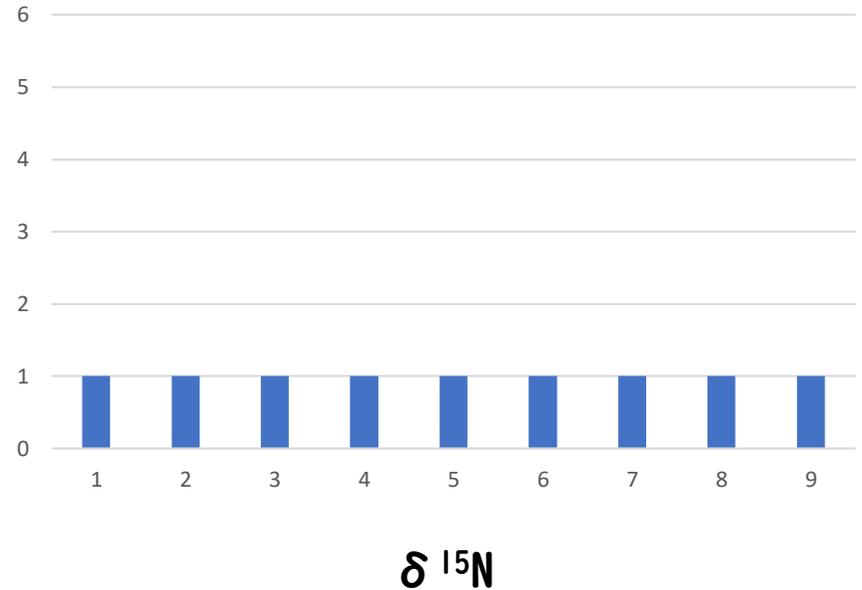
その数字を正しく解釈するために、  
ばらつきの値を教えてくださいませんか？

結果を正しく解釈するために  
不可欠なツール

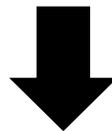
## 不確かさ小



## 不確かさ大



不確かさが小さい



測定値が真の値により近い可能性をもつ

不確かさを評価するために

**標準偏差**が必要

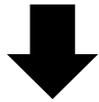
不確かさを評価するために

標準偏差が必要

データのばらつきの大きさの程度

とある母集団or試料（標本）において

1.平均値を求める



2.偏差を求める



3.分散を求める



4.標準偏差を求める

とある母集団or試料（標本）において

1. 平均値を求める



2. 偏差を求める



3. 分散を求める



4. 標準偏差を求める

※これらは統計学の基本なので  
多くの文献に記載されています。  
ネットにもたくさん転がっています。

標準偏差|

検索



とある**母集団**or**試料**（標本）において

1.平均値を求める



2.偏差を求める



3.分散を求める



4.標準偏差を求める

**母集団**とは？

**試料**とは？

JIS K 0211 「分析化学用語（基礎部門）」によると

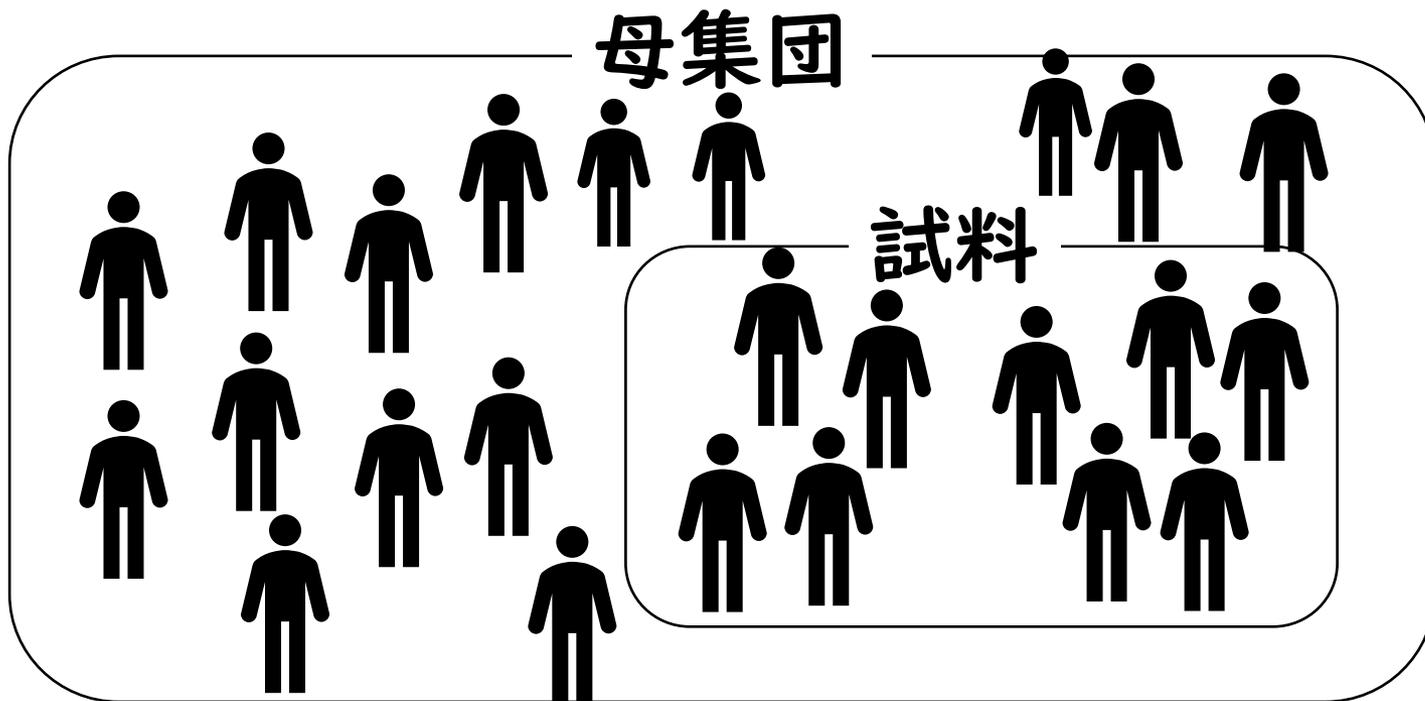
## 母集団（population）

調査の対象となる特性をもつすべてのものの集団

## 試料（sample）

母集団からその特性をもった、  
試験、検査、化学分析に供される物質

理想的には、母集団すべてを調べたいのだが、  
数が膨大すぎるので、現実的には難しいため、  
試料を調べ、統計学的に母集団を評価する。



調査項目

小学生の将来の夢

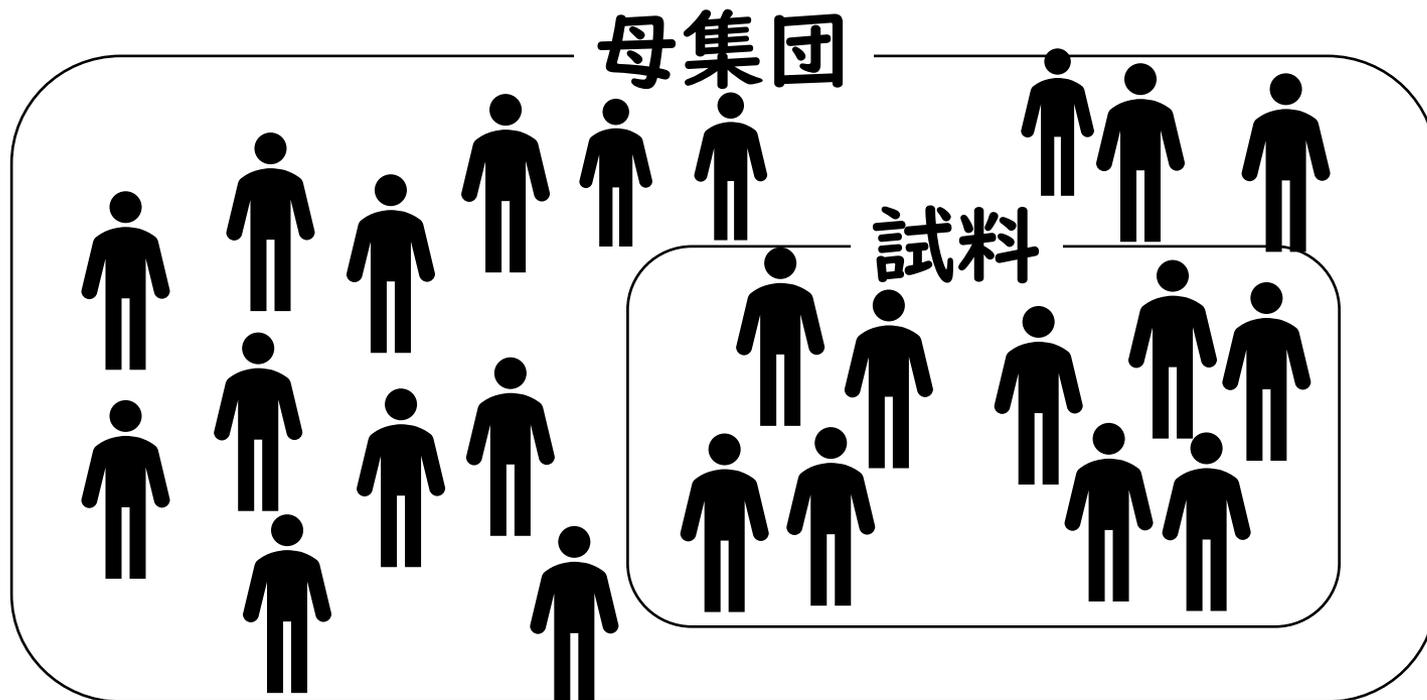
母集団

全国の小学生

試料 (標本)

アンケートに答えた小学生

# 母集団と試料における平均値と標準偏差



	母集団		試料 (標本)
平均値	母平均	導出	標本平均
標準偏差	母標準偏差		標本標準偏差

試料が少ないなど、母集団が存在しない場合もある。

## 例題

ジェンキンス研における  
研究室ゼミの発表時間の標準偏差を求めなさい

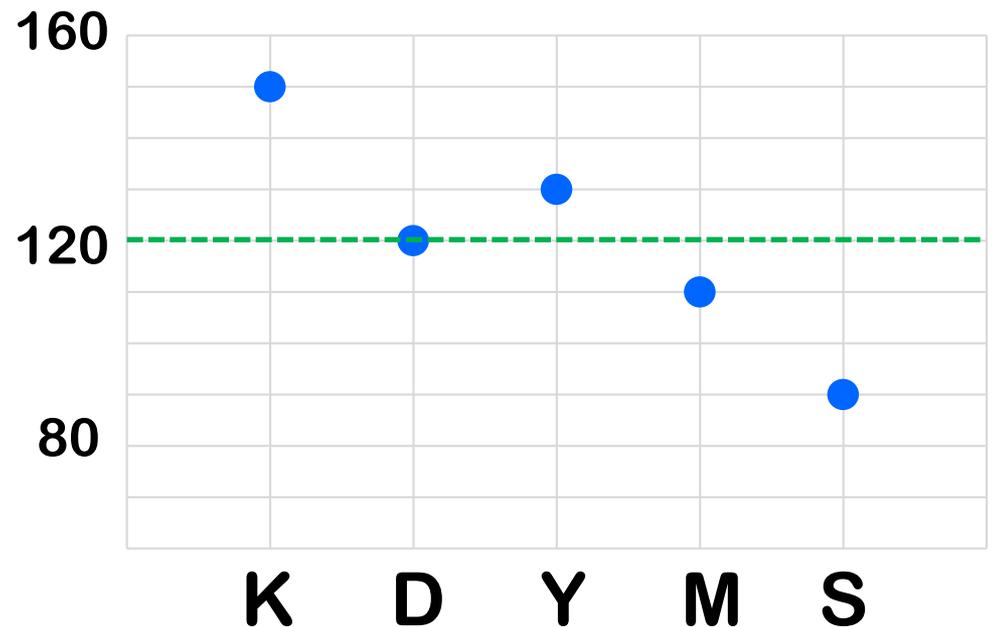
発表者	発表時間
K氏	150
D氏	120
Y氏	130
M氏	110
S氏	90

# 1. 平均値を求めよう

$$\begin{aligned} \text{平均値} &= \frac{1}{5} \sum_{t=1}^5 t_i \\ &= 120 \end{aligned}$$

発表者	発表時間(t)
K氏	150(t1)
D氏	120(t2)
Y氏	130(t3)
M氏	110(t4)
S氏	90(t5)

研究室ゼミの  
平均発表時間は**120分**



## 2. 偏差を求めよう

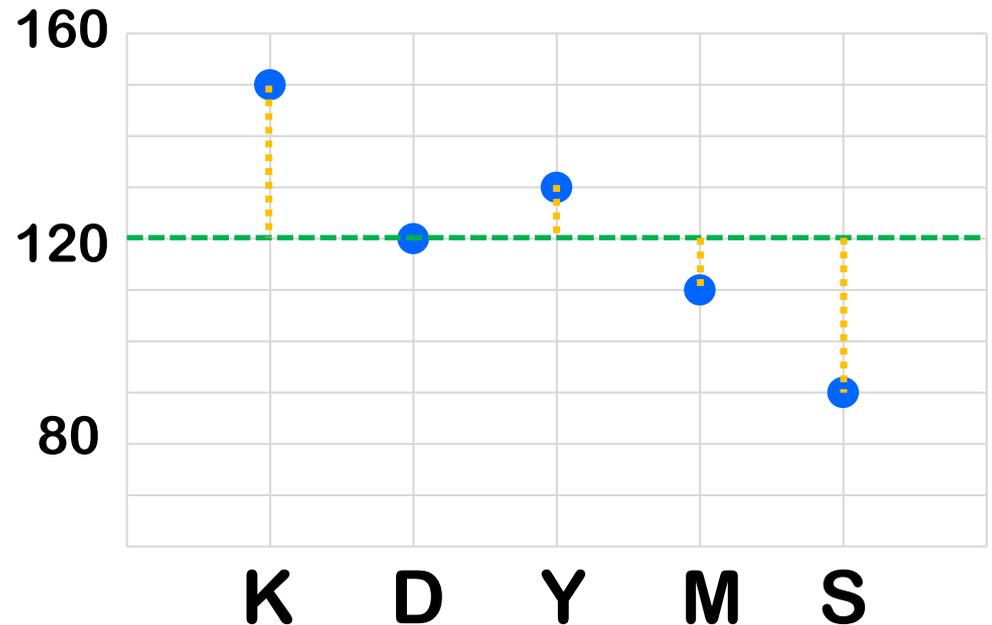


偏差とは  
各データと  
平均値の差



平均値から  
どのくらい  
離れているか

発表者	発表時間(t)	偏差
K氏	150(t1)	$150 - 120 = 30$
D氏	120(t2)	$120 - 120 = 0$
Y氏	130(t3)	$130 - 120 = 10$
M氏	110(t4)	$110 - 120 = -10$
S氏	90(t5)	$90 - 120 = -30$

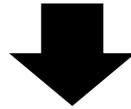


## 2. 偏差を求めよう

---



偏差を求めることで、  
ある値（学生の発表時間）がどのくらいズれているのかわかる。



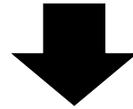
全体ではどのくらいズれているのか？

## 2.偏差を求めよう

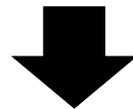
---



偏差を求めることで、  
ある値（学生の発表時間）がどのくらいズれているのかわかる。



全体ではどのくらいズれているのか？



全員のズレ（偏差）の平均を調べよう

## 2. 偏差を求めよう



発表者	発表時間(t)	偏差
K氏	150(t1)	$150 - 120 = 30$
D氏	120(t2)	$120 - 120 = 0$
Y氏	130(t3)	$130 - 120 = 10$
M氏	110(t4)	$110 - 120 = -10$
S氏	90(t5)	$90 - 120 = -30$

ジェンキンス研の学生の発表時間の平均値を求めると

$$\begin{aligned}(\text{平均値}) &= \frac{30 + 0 + 10 + (-10) + (-30)}{5} \\ &= 0\end{aligned}$$

0 (当たり前だけど)   
になっちゃった

全員のズレの平均がない (0) って直感的におかしい

0 (当たり前だけど)   
になっちゃった

全員のズレの平均がない (0) って直感的におかしい

そこで!!!!

平均偏差マン

参戦!!!



分散マン

参戦!!!

もうちょっとクオリティを上げたかった・・・

平均偏差・・・偏差の絶対値の平均

$$\frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$$

絶対値をとることで、  
平均からの“距離”を求め、  
その平均値を求める。

直感的にわかりやすいのだが、あまり利用されない。

- ・ **正と負が混合**していなければ、使える。
- ・ 絶対値は数学的に使いにくい

「場合分け」とか。詳しくは私も理解できておりません

# 3.分散を求めよう

分散・・・偏差の2乗和を自由度で除したもの

$$\frac{1}{n} \sum_{i=1}^n (x_i - \text{xave})^2 \text{ or } \frac{1}{n-1} \sum_{i=1}^n (x_i - \text{xave})^2$$

偏差を2乗することで、

- ・ 符号をすべて正にできる。
- ・ 数学的に扱いやすい

分散を利用することで、同じ条件でばらつきを比較できる

※自由度はまだ理解できていません！！！！勉強中です

### 3.分散を求めよう

$$\begin{aligned} \text{分散} &= \frac{1}{5} \sum_{i=1}^5 (t_i - \text{tave})^2 \\ &= 400 \end{aligned}$$

発表者	発表時間(t)
K氏	150(t1)
D氏	120(t2)
Y氏	130(t3)
M氏	110(t4)
S氏	90(t5)

### 3.分散を求めよう

$$\text{分散} = \frac{1}{5} \sum_{i=1}^5 (t_i - \text{tave})^2$$

$$= 400$$

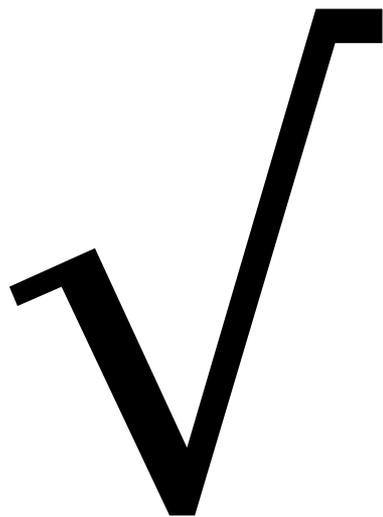
発表者	発表時間(t)
K氏	150(t1)
D氏	120(t2)
Y氏	130(t3)
M氏	110(t4)
S氏	90(t5)

しかし！！！！

“400” という数字は2乗しているので、元の単位（分）と合わない！！！！

単位を元に戻すために

# 平方根



を使う

# 4.標準偏差を求めよう

$$\begin{aligned}\sqrt{\text{分散}} &= \sqrt{\frac{1}{5} \sum_{i=1}^5 (t_i - t_{ave})^2} \\ &= \sqrt{400} \\ &= 20\end{aligned}$$

発表者	発表時間(t)
K氏	150(t1)
D氏	120(t2)
Y氏	130(t3)
M氏	110(t4)
S氏	90(t5)

# 4.標準偏差を求めよう

$$\sqrt{\text{分散}} = \sqrt{\frac{1}{5} \sum_{i=1}^5 (t_i - t_{ave})^2}$$

$$= \sqrt{400}$$

$$= 20$$

発表者	発表時間(t)
K氏	150(t1)
D氏	120(t2)
Y氏	130(t3)
M氏	110(t4)
S氏	90(t5)

これが**標準偏差**

# ジェンキンス研のゼミの発表時間



発表者	発表時間(t)
K氏	150(t1)
D氏	120(t2)
Y氏	130(t3)
M氏	110(t4)
S氏	90(t5)

平均値 120    ➡    平均120分のセミナーである

標準偏差 20    ➡    ±20分の時間差がある

**D, Y, Mは標準的, Kは長い, Sは早い**

**※データ数が多くなれば, 信頼度も上がる**

Excelの「STDEV」 → 「標本標準偏差 “s”」  
「STDEP」 → 「母標準偏差 “σ”」

扱うデータによっては，母集団の概念がなくなる

今回のケースでは

ジェンキンス研だけの話



母集団なし

$$\sigma = s$$

今回のデータを元に  
全国の研究室の話



母集団あり

$$\sigma = \frac{s}{\sqrt{n-1}}$$

- 標準偏差を求める流れ  
平均値 → 偏差 → 分散 → 標準偏差
- 平均値に意味をもたせるために  
標準偏差は必要
- ジェンキンス研における  
ゼミの発表時間は120分 ± 20分

うまくまとまってなくてごめんなさい。  
私も現在進行形で勉強しています。

“数値の意味”を考える  
きっかけになればうれしいです。

せっかく出したデータを

**無駄にしないで！！！！**