

**速習！情報 ～共通テスト対策講座～**

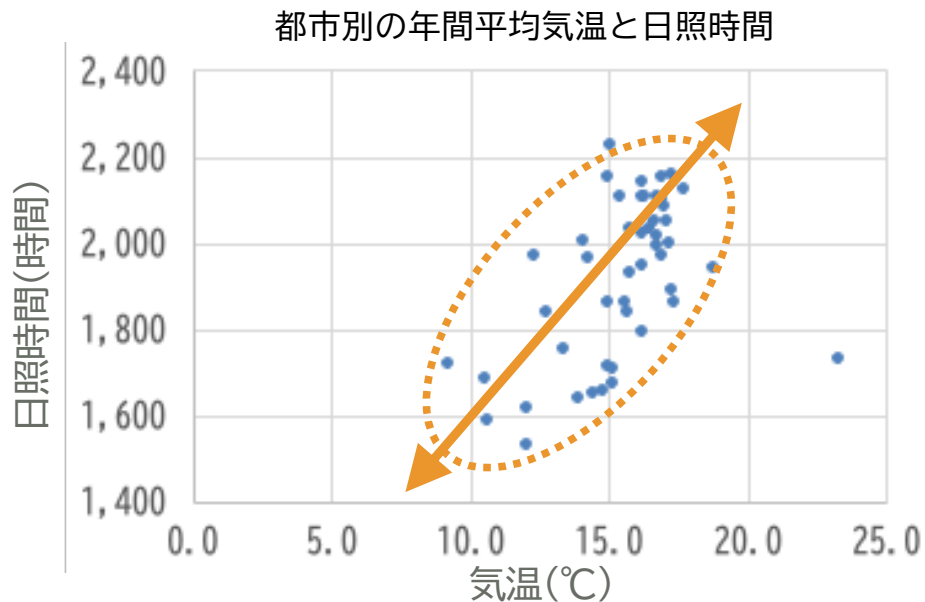
# **データ分析手法 (散布図と相関係数)**

**散布図、相関係数**

## (復習)散布図は、2項目のデータの分布、2項目間の相関の強さを示す

### 散布図

2項目のデータの分布、2項目間の相関の強さを示すのに適する。



大まかに  
「日照時間が増えれば気温が上がる」  
という関係性が読み取れる

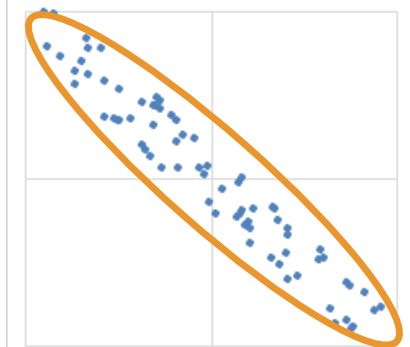
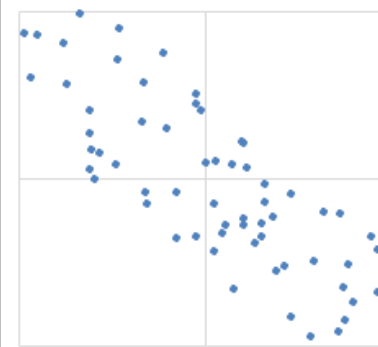
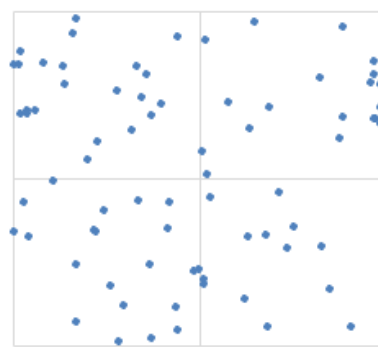
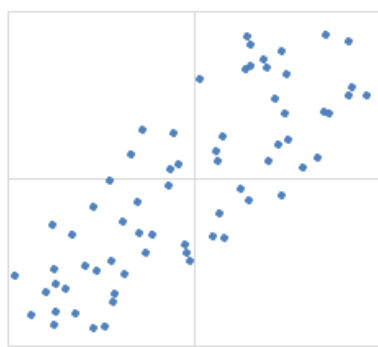
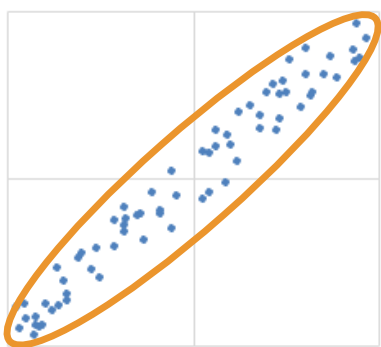
## 散布図から、2項目の間の相関関係の強さを読み取る

**相関関係**：2項目における、一方の項目の増減と他方の項目の増減が直線的な関係。  
このとき、2つの項目は**相関がある**という。

正の相関関係がある = 傾きが正

相関関係がない

負の相関関係がある = 傾きが負



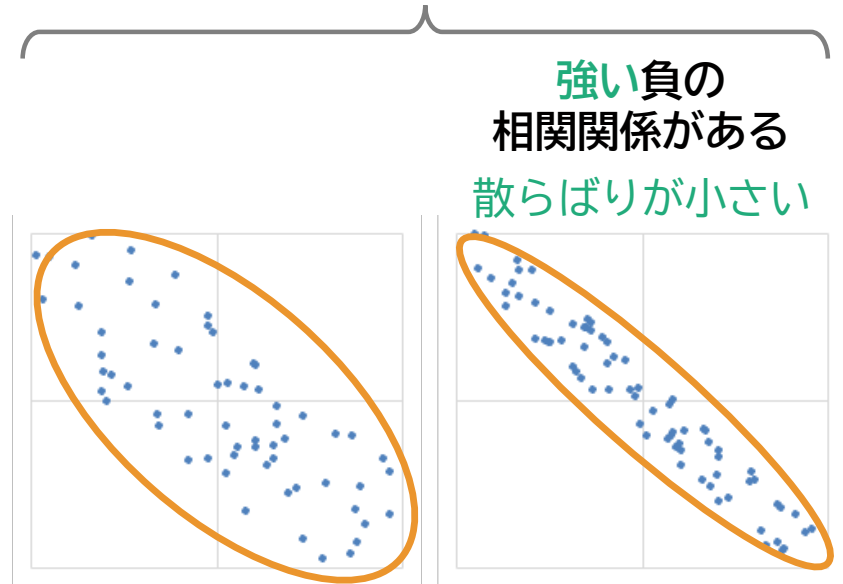
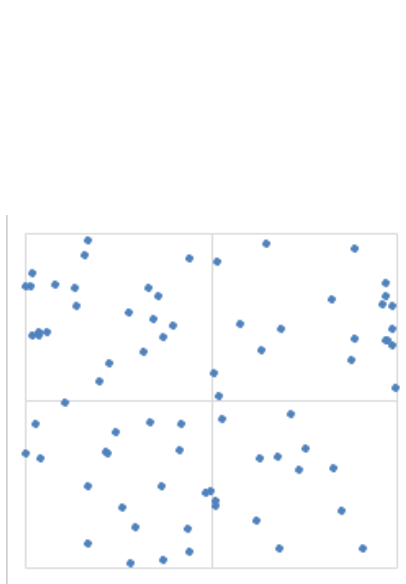
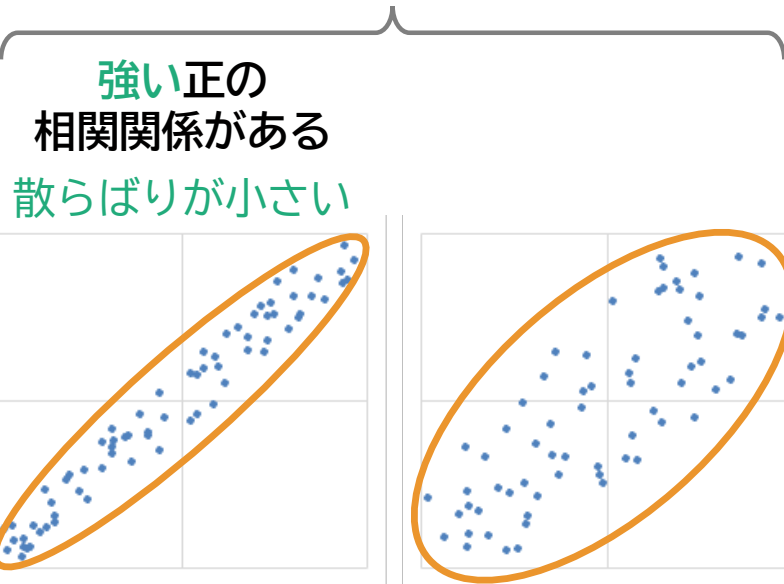
## 散布図から、2項目の間の相関関係の強さを読み取る

**相関関係**：2項目における、一方の項目の増減と他方の項目の増減が直線的な関係。  
このとき、2つの項目は**相関がある**という。

正の相関関係がある = 傾きが正

相関関係がない

負の相関関係がある = 傾きが負



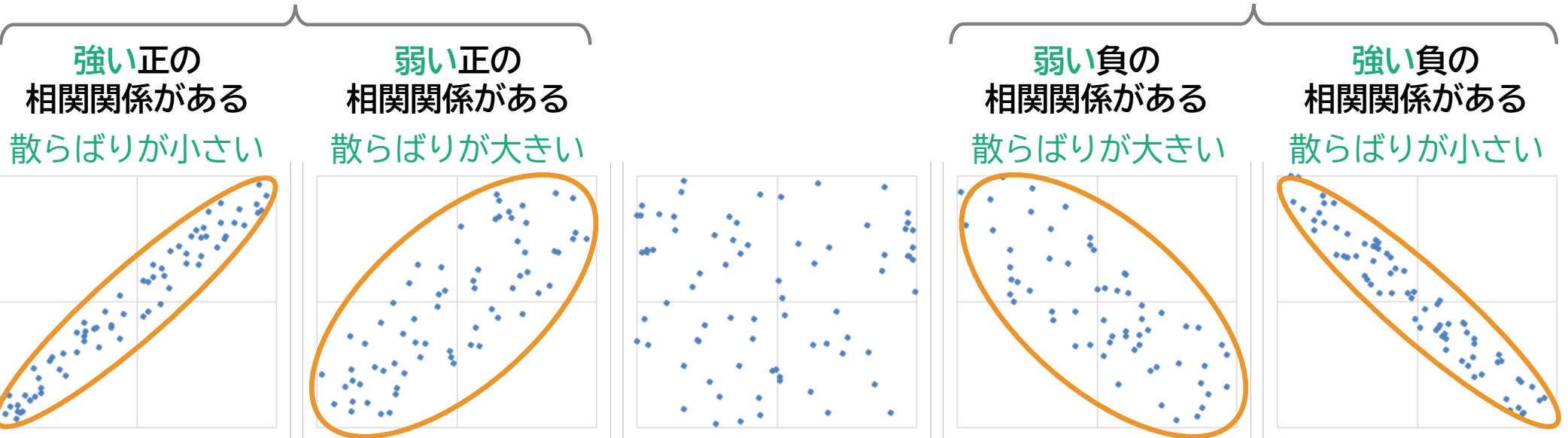
## 散布図から、2項目の間の相関関係の強さを読み取る

**相関関係**：2項目における、一方の項目の増減と他方の項目の増減が直線的な関係。  
このとき、2つの項目は**相関がある**という。

正の相関関係がある = 傾きが正

相関関係がない

負の相関関係がある = 傾きが負



## 散布図から、2項目の間の相関関係の強さを読み取る

**相関関係**：2項目における、一方の項目の増減と他方の項目の増減が直線的な関係。  
このとき、2つの項目は**相関がある**という。

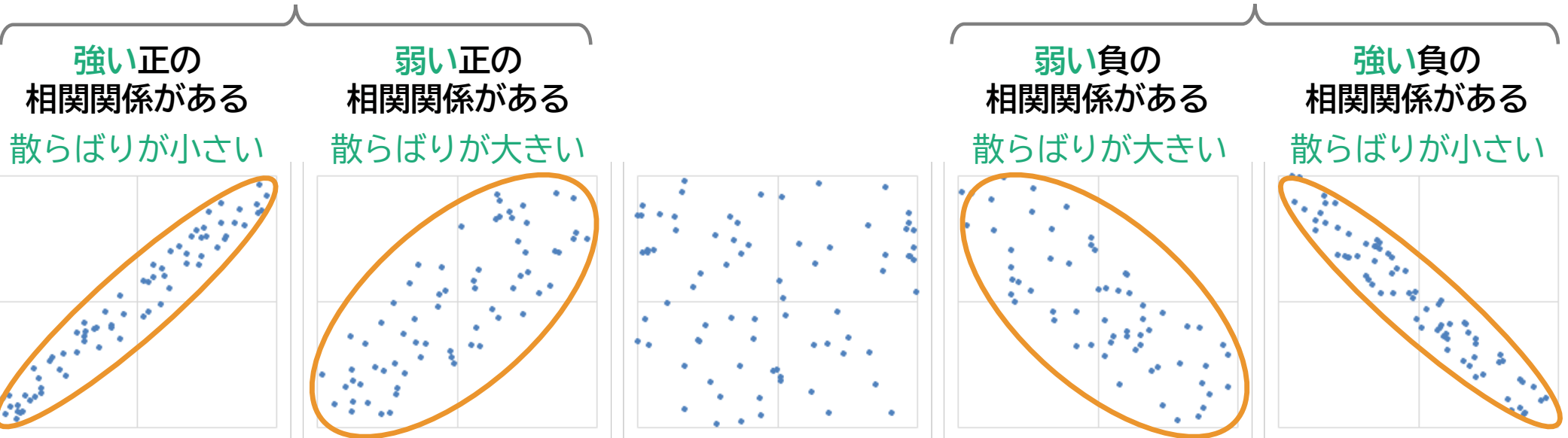
**相関係数** (相関関係を表す数値)

1 ← ————— 0 ————— → -1

**正の相関関係がある = 傾きが正**

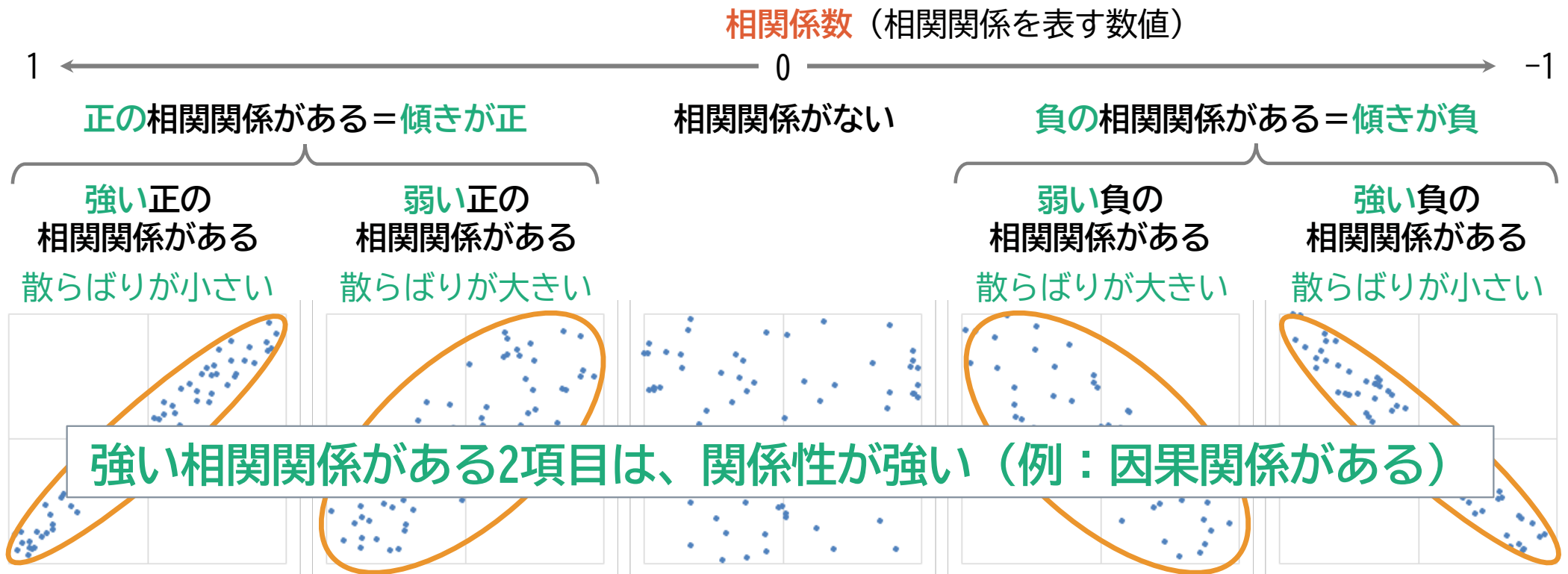
相関関係がない

**負の相関関係がある = 傾きが負**

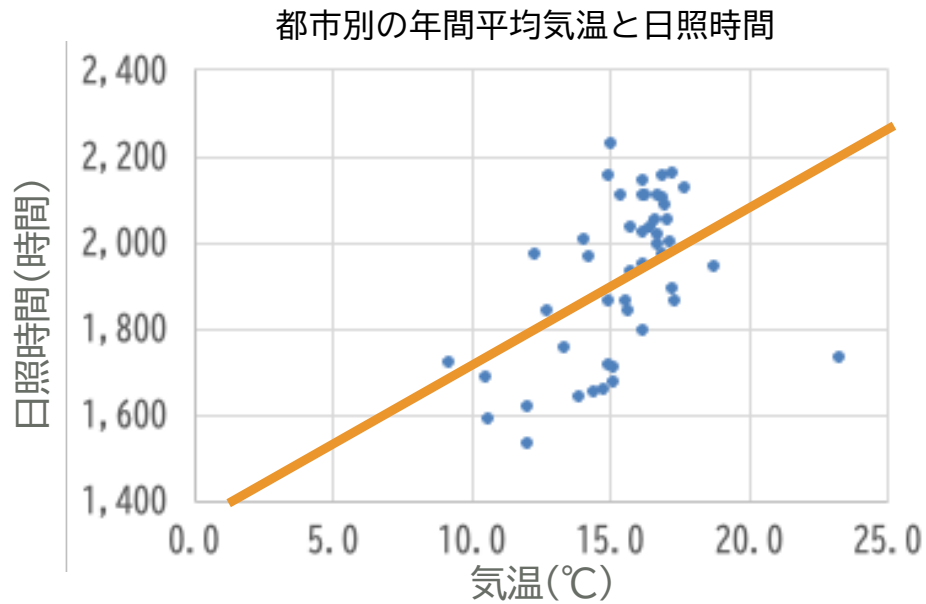


## 散布図から、2項目の間の相関関係の強さを読み取る

**相関関係**：2項目における、一方の項目の増減と他方の項目の増減が直線的な関係。  
このとき、2つの項目は**相関がある**という。



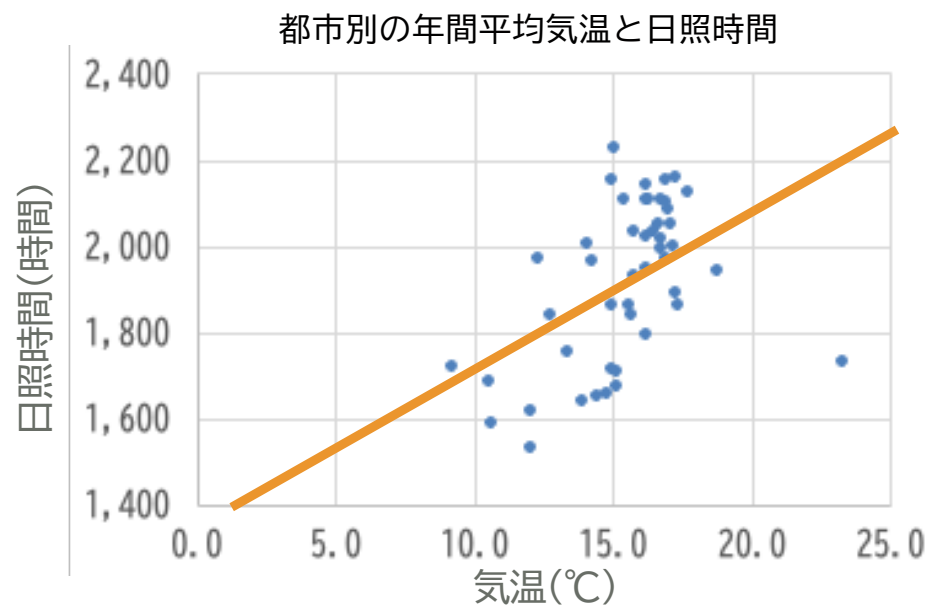
## 相関関係が強いときに因果関係がある場合もある



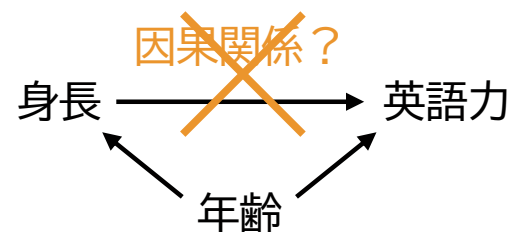
- 日照時間と気温は正の相関がある
  - 相関係数は0.46で、かなり相関がある
- 「日照時間が長いと気温が高くなる」という関係がありそう



## 相関関係が強くても因果関係があるとは限らず、交絡因子が存在する場合がある



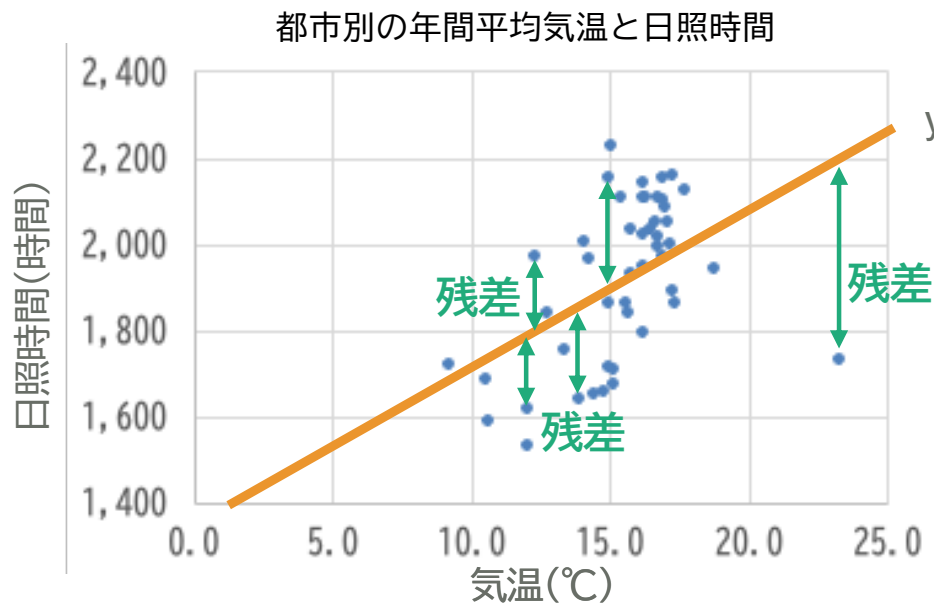
- 日照時間と気温は正の相関がある
  - 相関係数は0.46で、かなり相関がある
- 「日照時間が長いと気温が高くなる」という関係がありそう



身長と英語力に相関関係があった

両方に関係する要素(交絡因子)が原因かも

## 回帰直線とは、2項目の間関係性を直線の式で表したものの



2項目の間関係性を表す式を**回帰式**という。

回帰式を求めることを**回帰分析**という。

回帰式が直線の場合、**線形回帰(直線回帰、単回帰)**といい、求めた直線を**回帰直線**という。

(個々の値)と(回帰式のy座標)の差を**残差**といい、残差の2乗の合計が最小になるように回帰式を求める方法を**最小二乗法**という。

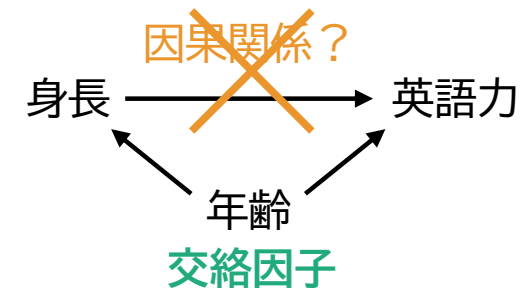
# 「データ分析手法(散布図と相関係数)」の要点



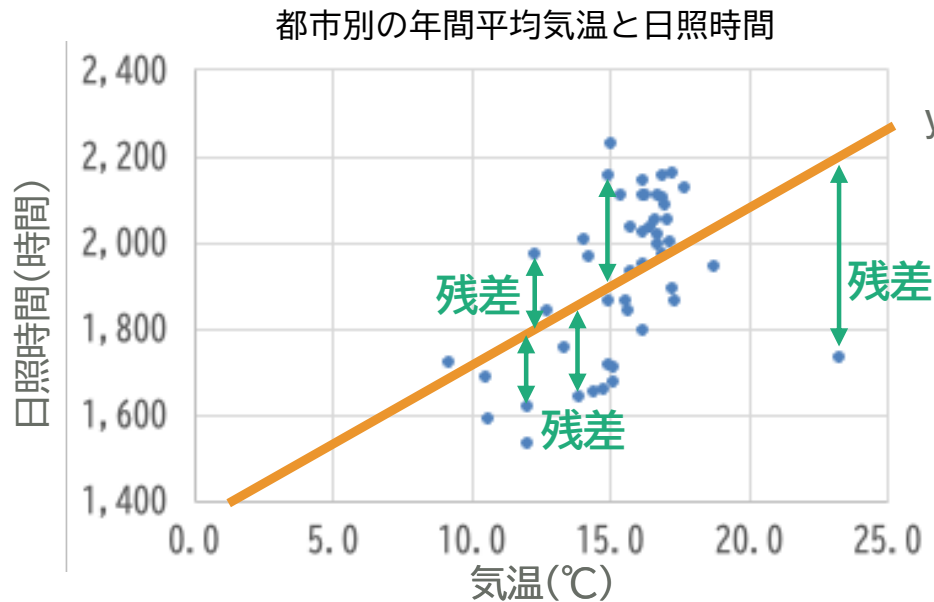
## 「データ分析手法(散布図と相関関係)」の要点2

相関関係が強くても因果関係があるとは限らない。交絡因子が存在する場合がある

- 日照時間と気温は正の相関がある
- 「日照時間が長いと気温が高くなる」という関係がありそう
- 身長と英語力に相関関係があった
- 両方に関係する要素(交絡因子)が原因かも



## 「データ分析手法(散布図と相関関係)」の要点3



2項目の間の関係性を表す式を**回帰式**という。

回帰式を求めることを**回帰分析**という。

回帰式が直線の場合、**線形回帰(直線回帰、単回帰)**といい、求めた直線を**回帰直線**という。

(個々の値)と(回帰式のy座標)の差を**残差**といい、残差の2乗の合計が最小になるように回帰式を求める方法を**最小二乗法**という。