

# 表計算ソフトによるデータ分析

年 組 氏名 \_\_\_\_\_

# 目次

<b>第Ⅰ部 操作関係</b>	<b>3</b>
1 表計算ソフトの基本操作	4
1.1 計算の基本 . . . . .	4
1.2 セル番地を用いて . . . . .	4
1.3 基本的な関数の使い方 . . . . .	5
1.4 繰り返し操作(コピー) . . . . .	6
1.5 相対参照と絶対参照 . . . . .	7
2 その他関数の使い方	8
2.1 数値計算 . . . . .	8
2.2 統計 . . . . .	8
2.3 論理式 . . . . .	9
2.4 まとめ . . . . .	9
3 グラフの描き方	10
3.1 グラフの種類 . . . . .	10
3.2 円グラフ . . . . .	11
3.3 棒グラフ, 散布図, その他 . . . . .	13
<b>第Ⅱ部 回帰分析</b>	<b>14</b>
4 回帰分析	15
4.1 回帰分析とは . . . . .	15
4.2 単回帰分析 . . . . .	15
4.3 最小誤差二乗法 . . . . .	15
4.4 絶対誤差二乗法 . . . . .	16
4.5 表計算ソフトでの分析 . . . . .	17
4.6 予測関数 . . . . .	19
<b>第Ⅲ部 その他の知識</b>	<b>20</b>
4.7 重回帰分析 . . . . .	20
4.8 過学習 . . . . .	21
4.9 サブタイトル . . . . .	22

# 第Ⅰ部

## 操作関係

# 1 表計算ソフトの基本操作

## 1.1 計算の基本

- 必ず半角で入力する.
- 計算するときは、「=」で始める.

	記号	記入例	結果
和	+	= 3 + 2	5
差	-	= 5 - 3	2
商	/	= 1/2	0.5
積	*	= 3 * 2	6
冪乗	^	= 3^2	9
ルート	SQRT( )	=SQRT(4)	2

- 計算に順序をつけたいときには、()を使えばいい.

記入例	結果
= 1 + 2 + 3 * (4 + 5)	30
= 3 / (2 + 8)	0.3

## 1.2 セル番地を用いて

	A	B	C	D
1				
2	3			
3			5	
4				

- A2 と C5 の和を求めたい → 「= A2 + C5」で OK.
- メリット 1：数値の変更をしても、計算結果に反映される.
- メリット 2：コピーや貼り付けで、同様操作の繰り返しをやりやすい.

### 1.3 基本的な関数の使い方

複数個のデータの計算を簡単にできる基本的な関数を使ってみよう。

	A	B	C	D
1	1	10	-3	2
2	3	0	1	2
3	-2	4	-5	2
4	4	4	4.2	3

上の表に対して、下のような式を入力すると、簡単に計算ができる。例えば、A1 から A3 と、B1 か

	関数	A1 から A4 の…	A1 から D1 の…
和	SUM	=SUM(A1:A4)	=SUM(A1:D1)
平均	AVERAGE	=AVERAGE(A1:A4)	=AVERAGE(A1:D1)
最大値	MAX	=MAX(A1:A4)	=MAX(A1:D1)
最小値	MIN	=MIN(A1:A4)	=MIN(A1:D1)
中央値	MEDIAN	=MEDIAN(A1:A4)	=MEDIAN(A1:D1)

ら B4 までの和を求めたいときは、

$$=SUM(A1:A3)+SUM(B1:B4) \quad \text{または} \quad =SUM(A1:A3, B1:B4)$$

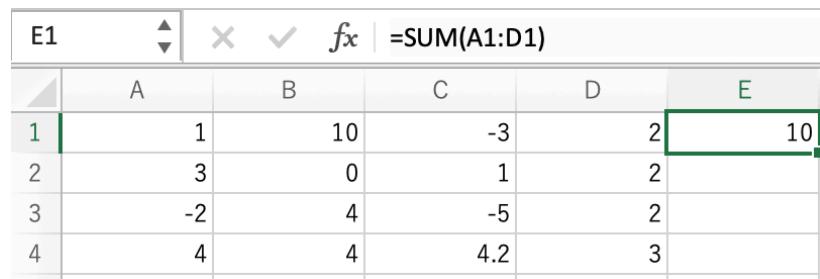
と書けばいい。

## 1.4 繰り返し操作(コピー)

	A	B	C	D	E
1	1	10	-3	2	=SUM(A1:D1)
2	3	0	1	2	
3	-2	4	-5	2	
4	4	4	4.2	3	

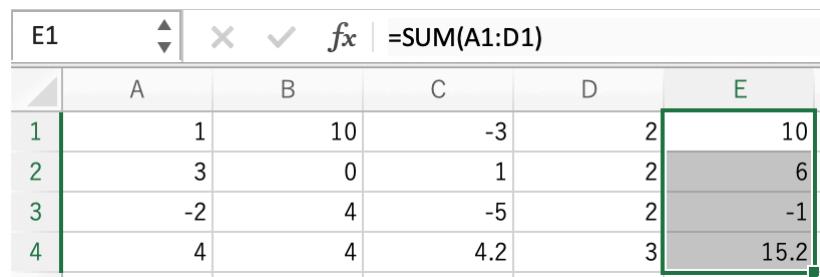
各行の和を計算したいときに、毎回「=SUM(… )」と書いていると大変である。そこで使えるのがコピーである。

1. 右下の四角にカーソルを合わせる。



	A	B	C	D	E
1	1	10	-3	2	10
2	3	0	1	2	
3	-2	4	-5	2	
4	4	4	4.2	3	

2. マウスの左ボタン(マウスピッド)を長押ししながら、コピーする場所まで下ろして完成。



	A	B	C	D	E
1	1	10	-3	2	10
2	3	0	1	2	6
3	-2	4	-5	2	-1
4	4	4	4.2	3	15.2

この操作は、「=A1+B1」などの四則演算でも同様にできる。

また、A5 に「=SUM(A1:A4)」と書き、右にコピーすることもできる。

「E1 の和の計算を E3 にだけ適用したい」など、部分的にコピペしたい時には、

E1 のセルを選択し、Ctrl + C でコピー → E3 のセルを選択し、Ctrl + V でペースト

でもできる。

## 1.5 相対参照と絶対参照

前ページでコピーについて学んだ.

例えば、全ページで学んだ方法で E1 を下へコピーすれば、「A2+B2」となる。また、A5 を右へコピーすれば、「=SUM(B1 : B4)」になる、これを相対参照という。

	A	B	C	D	E
1	1	10	-3	2	=A1+B1
2	3	0	1	2	
3	-2	4	-5	2	
4	4	4	4.2	3	
5	=SUM(A1 : A4)				

例えば、E1 を下へコピーするが、A1 を固定して計算したい場合、つまり、A1+B2, A1+B3, … をしたいとき、E1 には「=A\$1+B1」と書く。行、列の記号・番号の前に \$ マークを付けることで、固定してコピーすることが可能になる。このように、参照する場所を指定する方法を絶対参照という。

絶対参照の方法がメリットが大きい計算として、偏差を求めるときである。

	A	B	C
1		値	偏差
2	データ 1	2	=A2-A\$20
3	データ 2	4	
4	データ 3	1	
⋮	⋮		
19	データ 19	8	
20	平均	=AVERAGE(A2:A19)	

上の図のような 19 個のデータの偏差を調べる際に、\$ を付けることで、データから引く値「平均」を固定できる。

## 2 その他関数の使い方

### 2.1 数値計算

関数	説明	書き方	結果
INT(数値)	整数部分を求める	=INT(4.32)	4
MOD(数値, 除数)	数値を除数で割ったあまり	=MOD(43, 5)	3
ROUND(数値, 桁数)	桁数指定して数値を四捨五入	=ROUND(1.333, 2) =ROUND(1323, -2)	1.33 1300
RAND()	0 以上 1 未満の乱数を生成	=RAND()	

乱数は、Windows であれば Shift + F9 で再計算できる（らしい）が、Mac の場合この方法では再計算できない。Excel であれば、下図の右側にある「再計算実行」ボタンで再計算できる。



### 2.2 統計

関数	説明	書き方	結果
RANK(数値, 参照, [順序])	参照範囲での数値の順位を計算	=RANK(A1, A1:A5, 0) [順序]=0 で降順, 1 で昇順	3
VAR.P(数値)	標本分散を求める	=VAR.P(A1:A5)	
STDEV.P(数値)	母集団に対する標準偏差の計算	=STDEV.P(A1:A5)	
CORREL(配列 1, 配列 2)	配列 1 と配列 2 の相関係数の計算	=CORREL(A1:A5, B1:B5)	

母集団に対して VAR.P, STDEV.P で計算し、標本を対象にするときは VAR.S, STDEV.S で計算する。数学 1A で学んだものは、母集団に対する分散、標準偏差である。母集団や標本についての詳しい説明はまた別項目で。

## 2.3 論理式

関数	説明	書き方
IF(論理式, 真の場合, 偽の場合)	式が真か偽かで場合分け	=IF(A1>0, "正の数", 0)
AND(論理式 1, 論理式 2, …)	全ての論理式が真のとき TRUE 1つでも偽であれば FALSE	=AND(A1>0, A2<0)
OR(論理式 1, 論理式 2, …)	真である論理式があるとき TRUE 全て偽であれば FALSE	=OR(A1>0, A2<0)
NOT(論理式)	論理式が真であれば FALSE 偽であれば TRUE	=NOT(A1>0)

IF 関数については、文字を返したいときには” ”をつける。数式を返すときには不要。  
また、AND のことを論理積、OR のことを論理和、NOT のことを否定という。

関数	説明	書き方
COUNTIF(範囲, 検索条件)	検索条件を満たすものの個数を計算	=COUNTIF(A1:A5, ">0")
SUMIF(範囲, 検索条件)	検索条件を満たすものの合計を計算	=SUMIF(A1:A5, ">0")

## 2.4 まとめ

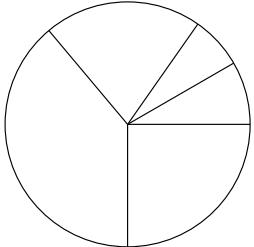
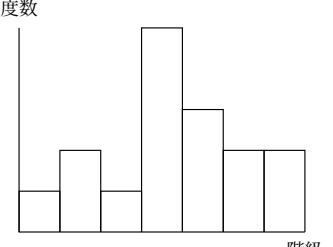
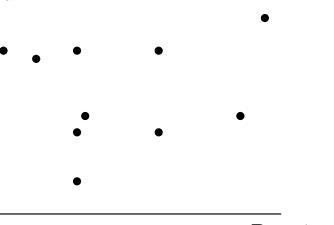
さまざまな計算方法と実際の計算結果を以下に載せておく。参考にしてみてほしい。

A	B	C	D	E	F	G	H	I	J
1 2.1				2.2					
2		記入例	結果		配列1	配列2		書き方	結果
3 整数部	1.43431	=INT(B3)	1		21	89	F3の配列1内順位(降順)	=RANK(F3,F3:F10,1)	3
4 剰余	234	=MOD(B4,10)	4		32	23	配列1の分散	=VAR.P(F3:F10)	483.984375
5 四捨五入	12.4567	=ROUND(B5,2)	12.46		12	64	配列1の標準偏差	=STDEV(F3:F10)	23.5186097
6	32145	=ROUND(B6,-2)	32000		46	1	相関係数	=CORREL(F3:F10,G3:G10)	-0.5614349
7 乱数		=RAND()	0.40421886		13	63			
8					85	27			
9					41	15			
10 2.3					35	74			
11	記入例	結果							
12 1	=IF(A12>0, SUM(A12:A17),0)	21							
13 2	=AND(A12>0,A13>0)	TRUE							
14 3	=OR(A12>1, A13>1, A14>1)	TRUE							
15 4	=NOT(A15>0)	FALSE							
16 5	=COUNTIF(A12:A17,">3")	3							
17 6	=SUMIF(A12:A17,">3")	15							

### 3 グラフの描き方

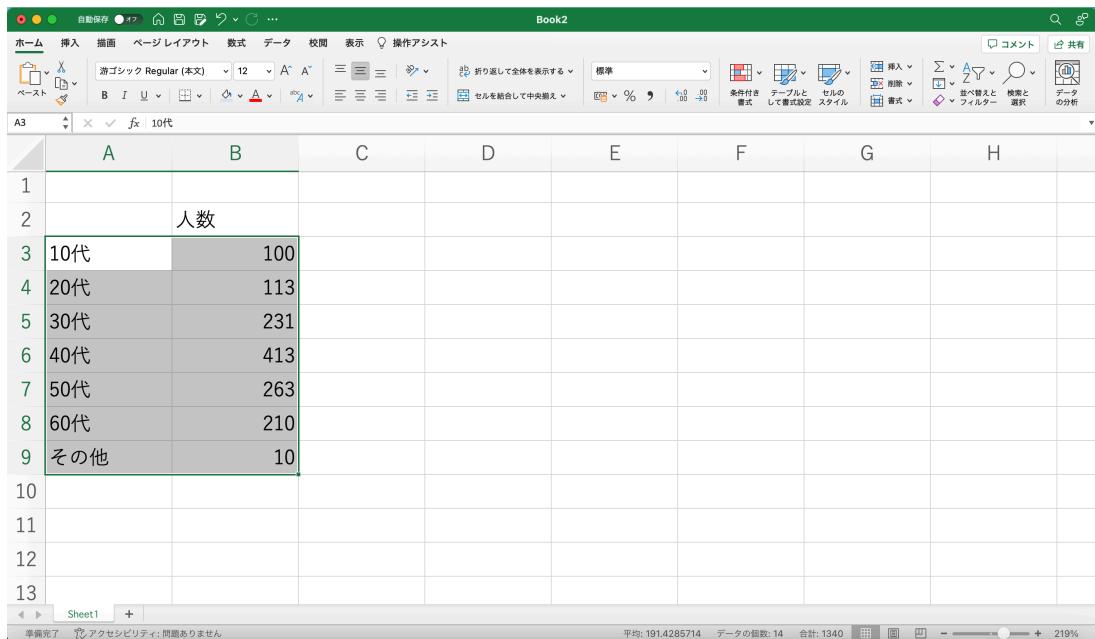
#### 3.1 グラフの種類

グラフにはさまざまな種類があるが、何を知りたいかという目的によって使うグラフを選ぶ必要がある。

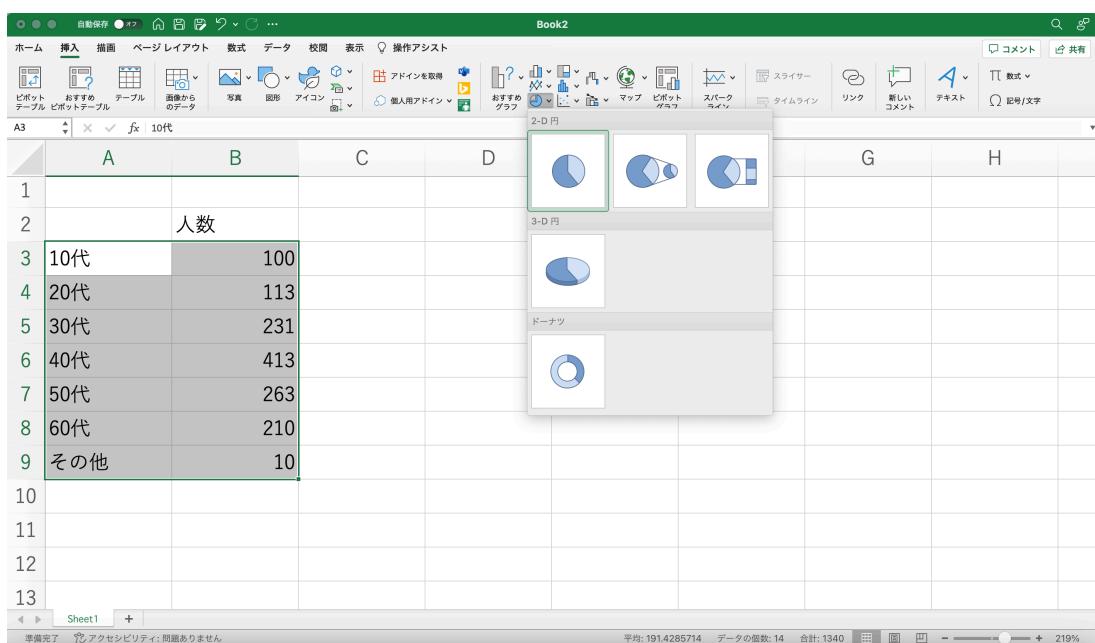
グラフの種類	円グラフ	棒グラフ	散布図
図			
特徴	<ul style="list-style-type: none"><li>・全体に対する割合が見やすい</li><li>・割合比較がしやすい</li><li>・</li></ul>	<ul style="list-style-type: none"><li>・度数差が見やすい</li><li>・分布が見やすい</li><li>・</li></ul>	<ul style="list-style-type: none"><li>・2つのデータの傾向を見れる</li><li>・</li><li>・</li></ul>

## 3.2 円グラフ

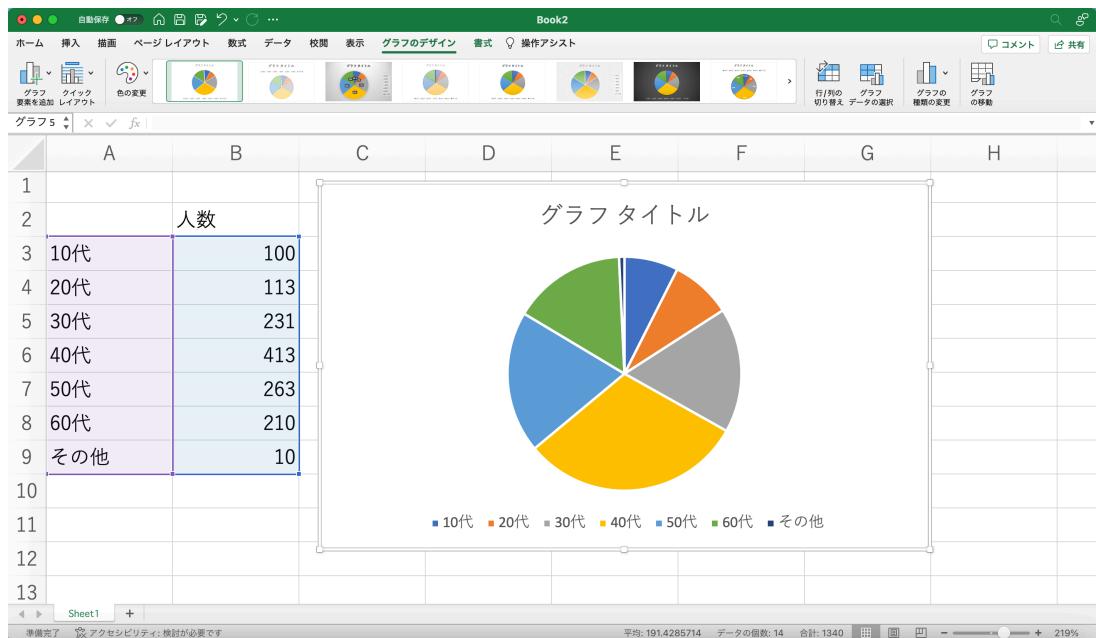
1. データの範囲を選択. 要素名も選択しておくといいかも.



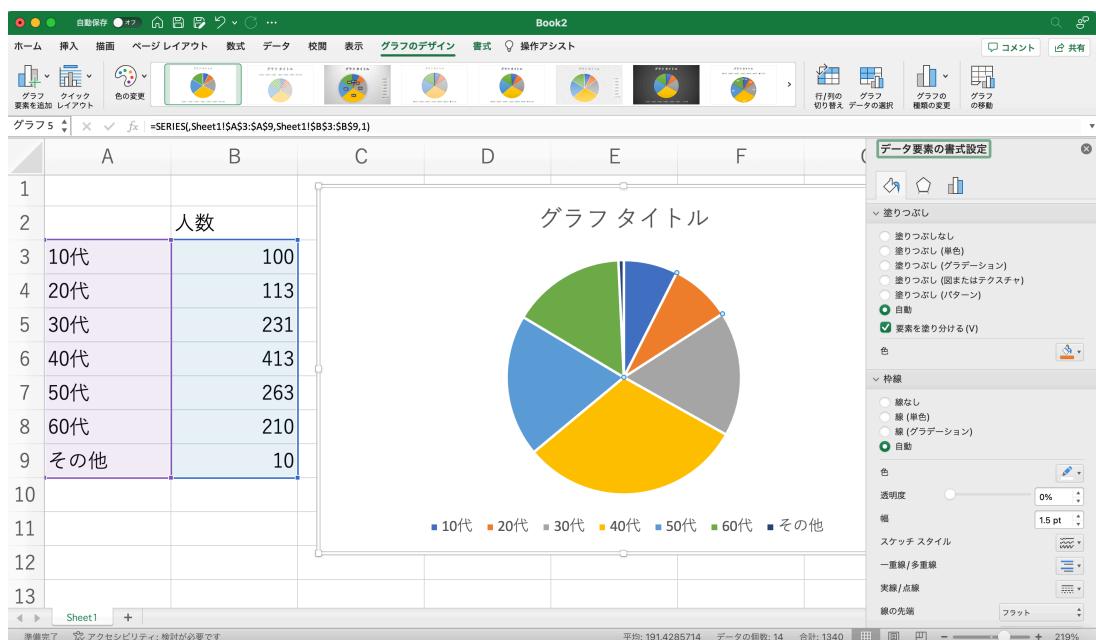
2. 「挿入」から「円グラフ」を選択. 形状を選択してクリック.



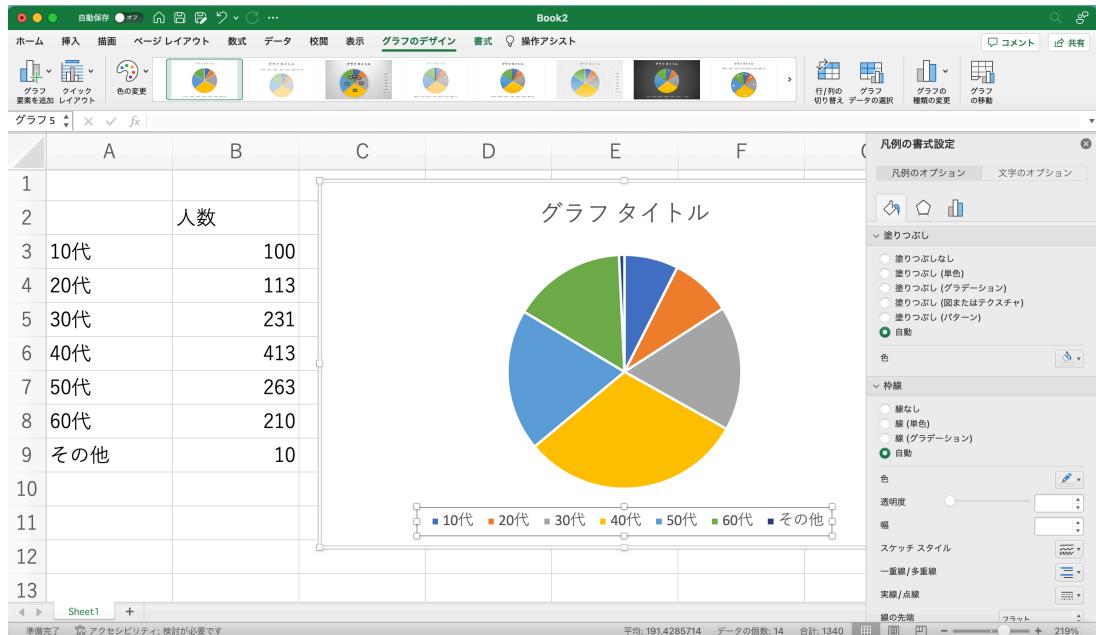
### 3. グラフが表示される。



### 4. 円をクリックすると、塗りつぶしや枠線などの設定ができる。



5. 凡例をクリックすれば、書式設定等ができる。



### 3.3 棒グラフ, 散布図, その他

棒グラフや円グラフなどの他のグラフも同じような手順で描くことができる。

## 第Ⅱ部 回帰分析

## 4 回帰分析

### 4.1 回帰分析とは

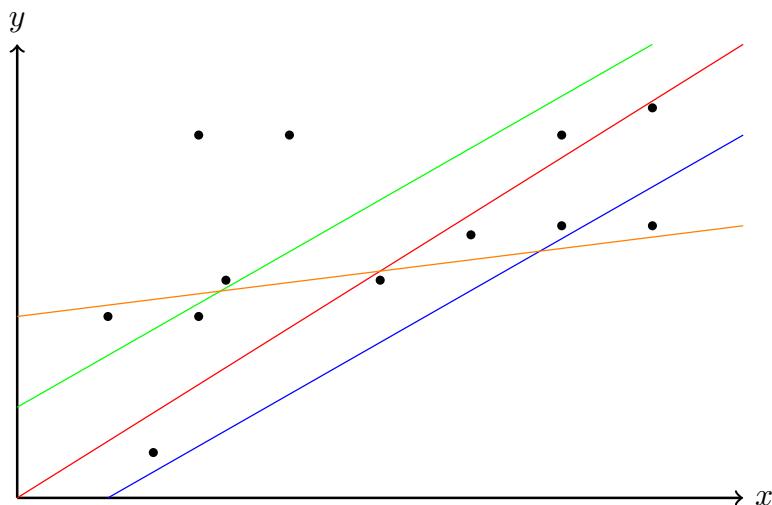
データの間の関係性を、式によって説明する分析手法である。例えば、気温によってアイスの売り上げを予測しようというものが回帰分析である。ここで、予測したいデータ（ここではアイスの売り上げ）のことを「目的変数」といい、予測したいデータの要因（ここでは気温）のことを「説明変数」という。

説明変数を  $x$ 、目的変数を  $y$  として、 $y = ax + b$  のように関係式を表し、説明変数を変化させることで目的変数を予測しようというのが、回帰分析の目標である。

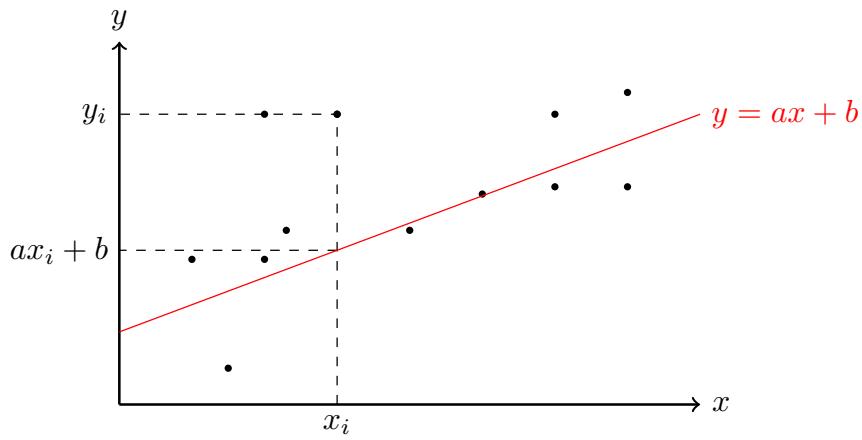
### 4.2 単回帰分析

先ほどの例であれば、「アイスの売り上げ」を「気温」で予測しようとした。このように、目的変数を 1 つの説明変数で分析しようとするものを「単回帰分析」という。

### 4.3 最小誤差二乗法



点がデータ、直線を予測したものとしたとき、どの直線が予測として最も適切だろうか。最も適切な予測直線を決める手法として、最小誤差二乗法がある。



上の図において,  $x_i$  での予測値との差は  $|y_i - (ax_i + b)|$  と書ける. この差を全てのデータについて調べて, それぞれ二乗して足したものを考える.

$$E = \sum_{i=1}^n (y_i - (ax_i + b))^2$$

この誤差の和  $E$  が最も小さくなるように  $y = ax + b$  を決定するものが最小誤差二乗法である.

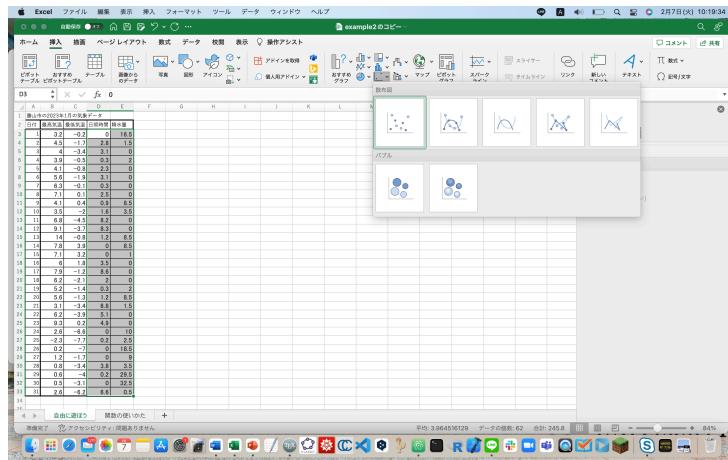
#### 4.4 絶対誤差二乗法

最小誤差二乗法での誤差を に変えたものを絶対誤差二乗法という.

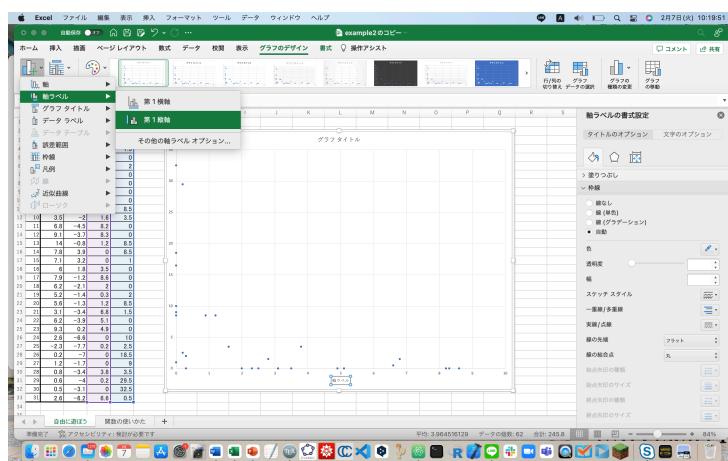
$$E = \sum_{i=1}^n |y_i - (ax_i + b)|$$

## 4.5 表計算ソフトでの分析

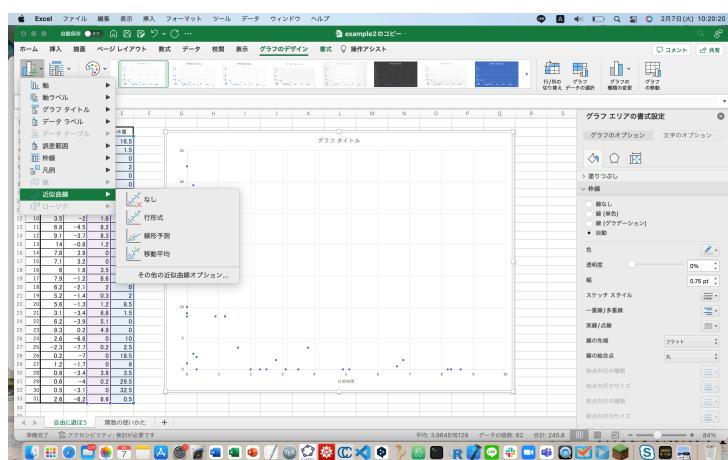
### 1. 散布図を描く。



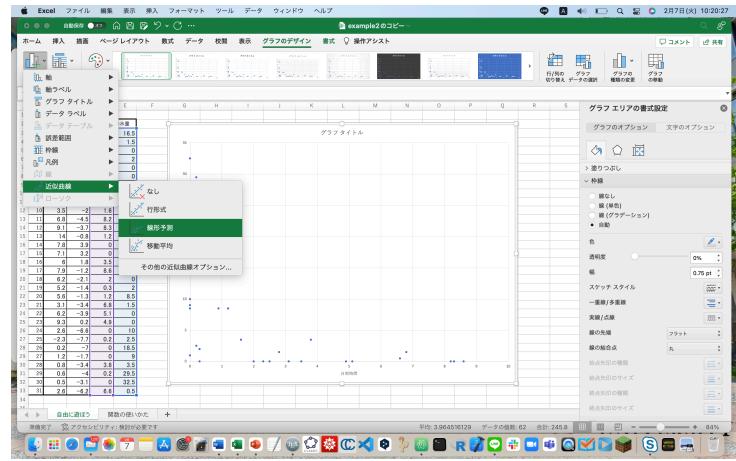
### 2. グラフの要素を追加から、軸名を追加できる。



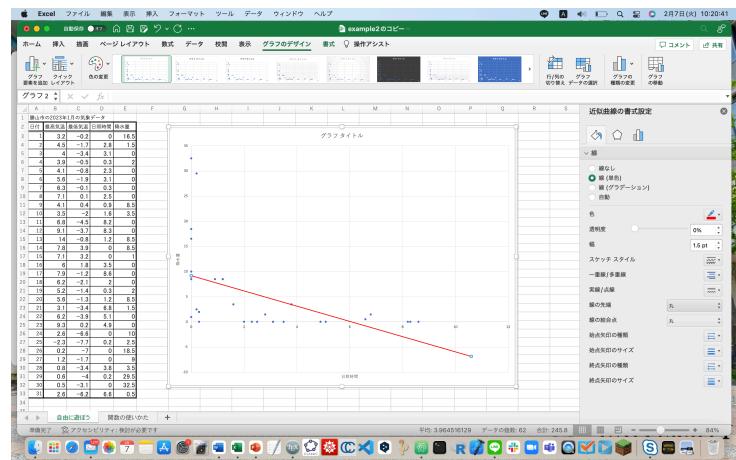
### 3. 同じようにグラフの要素を追加から、近似直線を追加できる。



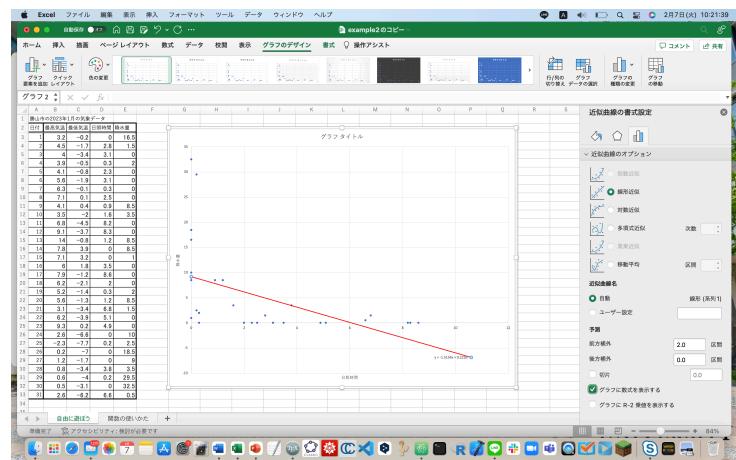
#### 4. ここでは、線形予測を選択。



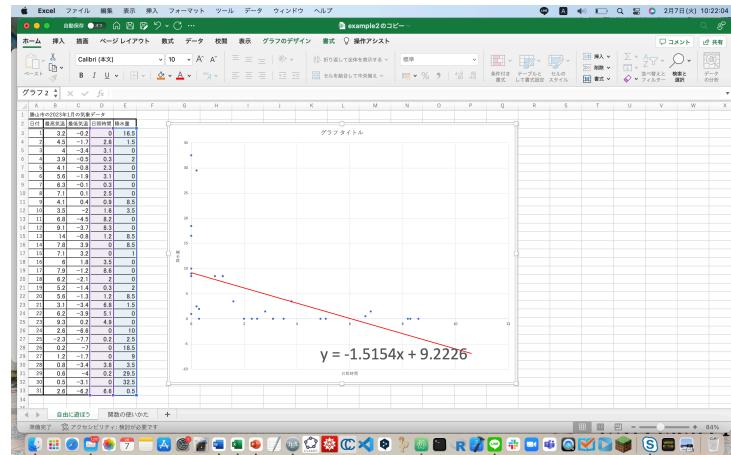
#### 5. 近似曲線が表示されるので、書式設定で自分の好きに設定。



#### 6. 書式設定から、グラフに数式を表示できる。



7. この数式の説明変数に数値を代入することで、予測値が算出できる。



## 4.6 予測関数

グラフを描いて予測を行ったが、これを関数1つで予測することもできる。

関数	書き方
FORECAST( $x$ の値, 既知の $y$ , 既知の $x$ )	=FORECAST(A51, A1:A50, B1:B50)

これで、既知のデータとして A1 から 50 に目的変数、B1 から 50 に説明変数が与えられている上での、A51 のときの予測値を算出することができる。算出方法は、グラフを描いて行ったものとほぼ同じである。

# 第 III 部

## その他の知識

### 4.7 重回帰分析

## 4.8 過学習

4.9 も