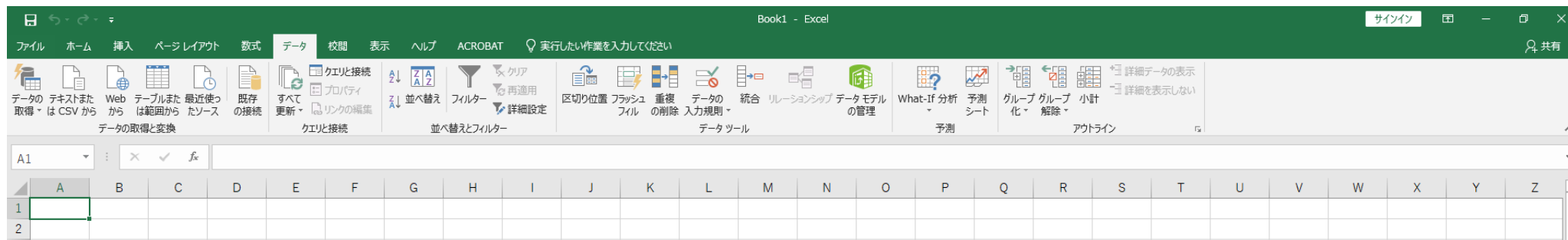


## 3.4 Excel 2019 による回帰分析

回帰分析が出来るように、Excel 2019 をセットアップする。

Excel 2013, Excel 2016 も同様の手順。

まず、エクセルの最初のページで、「データ」タブをマウスでクリックすると、下記の画面が出てくる。



様々な選択肢があり、一番左（「アウトライン」の右隣の右）に注目。

選択肢は何もない状態。

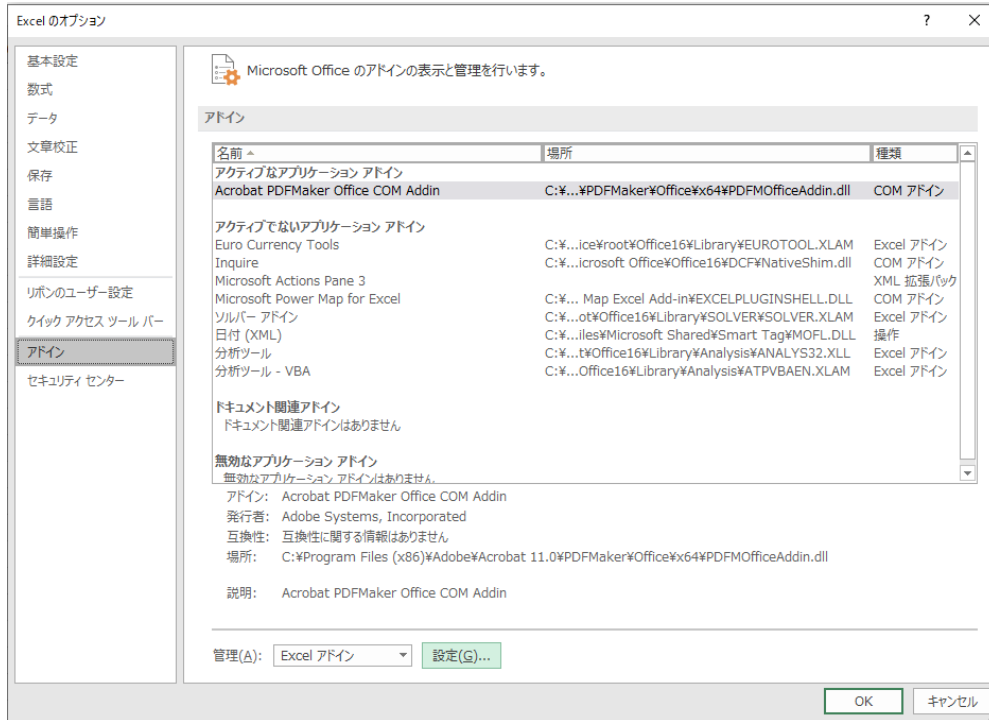
次に、「ファイル」タブをマウスでクリックすると下記の画面が現れる。



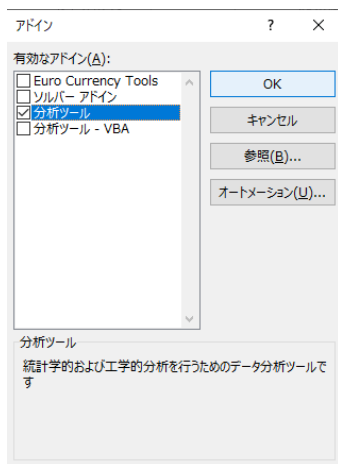
一番左下の「オプション」を選択して、下記の画面が出てくる。



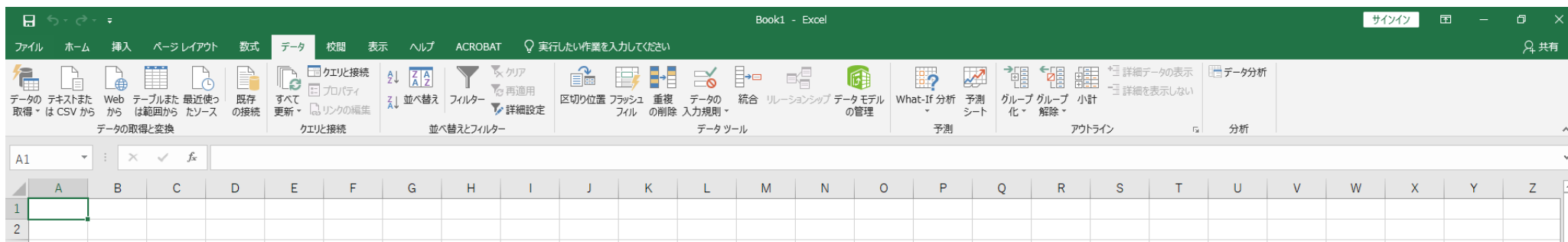
左の下から 2 番目の「アドイン」を選択して、下記の画面が出てくる。



「設定 (G)」 ボタンをクリックして、下記の画面が現れる。



「分析ツール」にチェックを入れて、「OK」ボタンをクリックすると、下記の「データ」タブの画面に戻る。



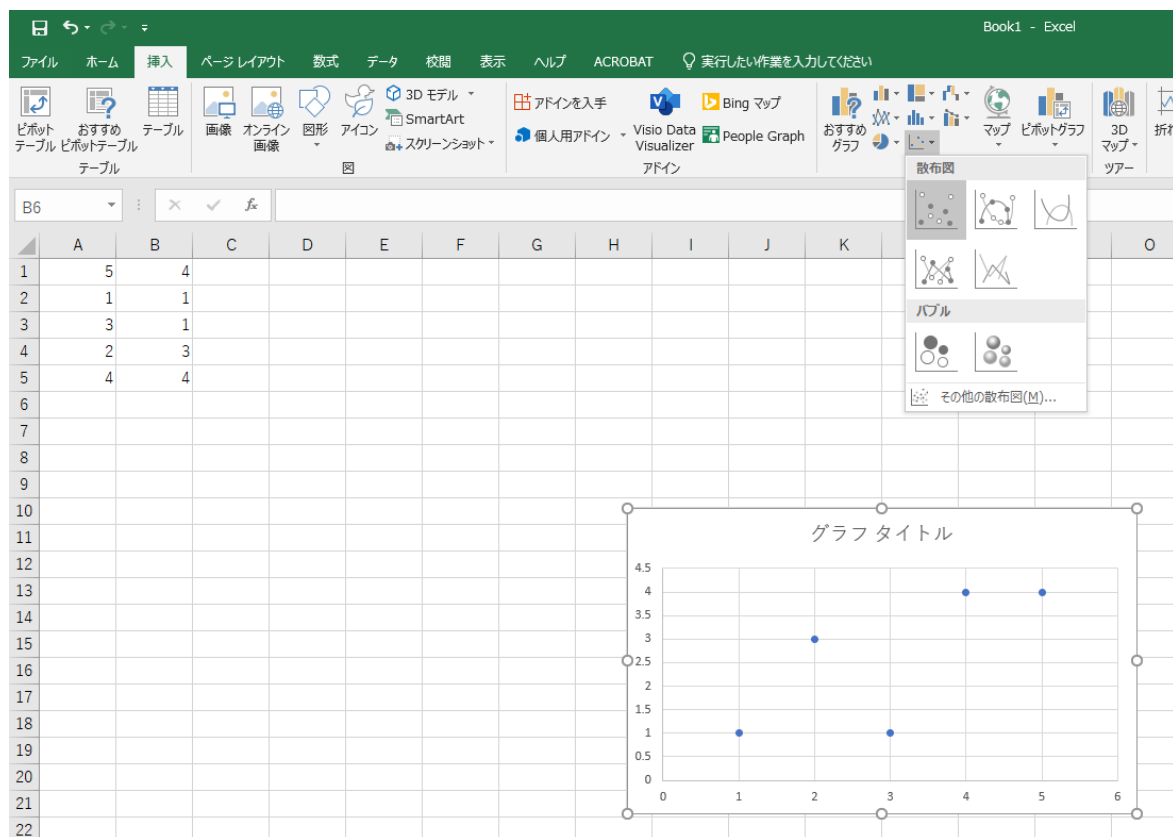
一番右に「データ分析」のタブが追加される。

これは一度だけ行えばよい。

次回からは自動的に「データ分析」のタブは追加されたままになる。

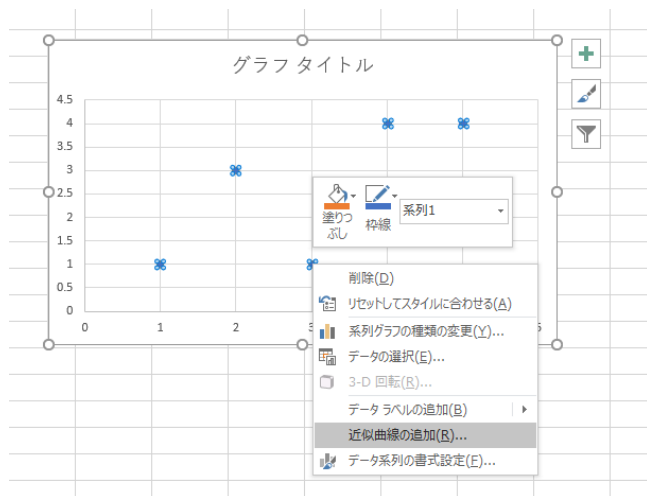


散布図を選び、さらに、左上の散布図を選ぶと、下記の画面が出る。



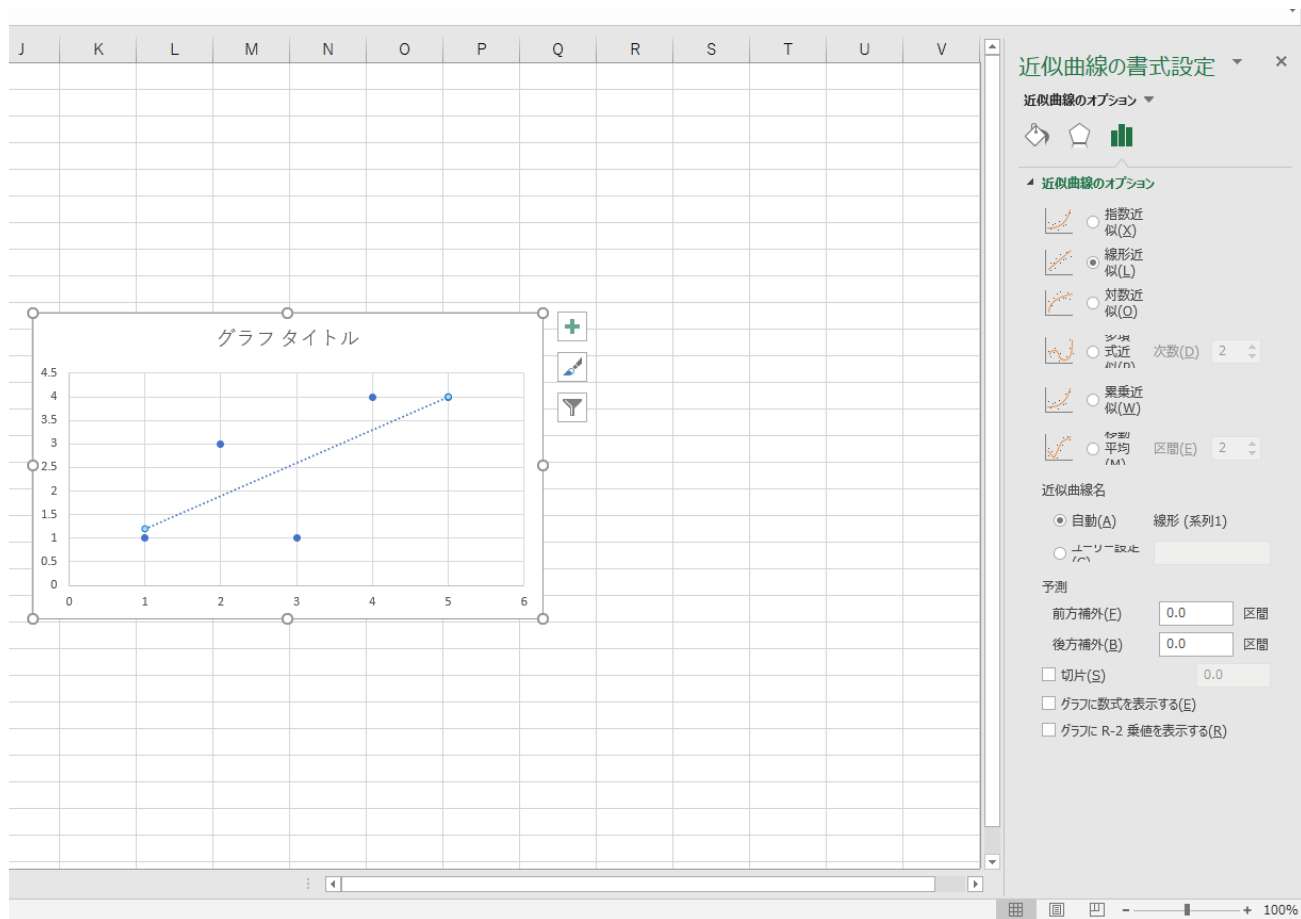
このように、横軸に A 列，縦軸に B 列の散布図が完成する。

グラフ内の 6 点のうちどれか一つをマウスで選び、マウスの右ボタンを押すと、下記の画面になる。

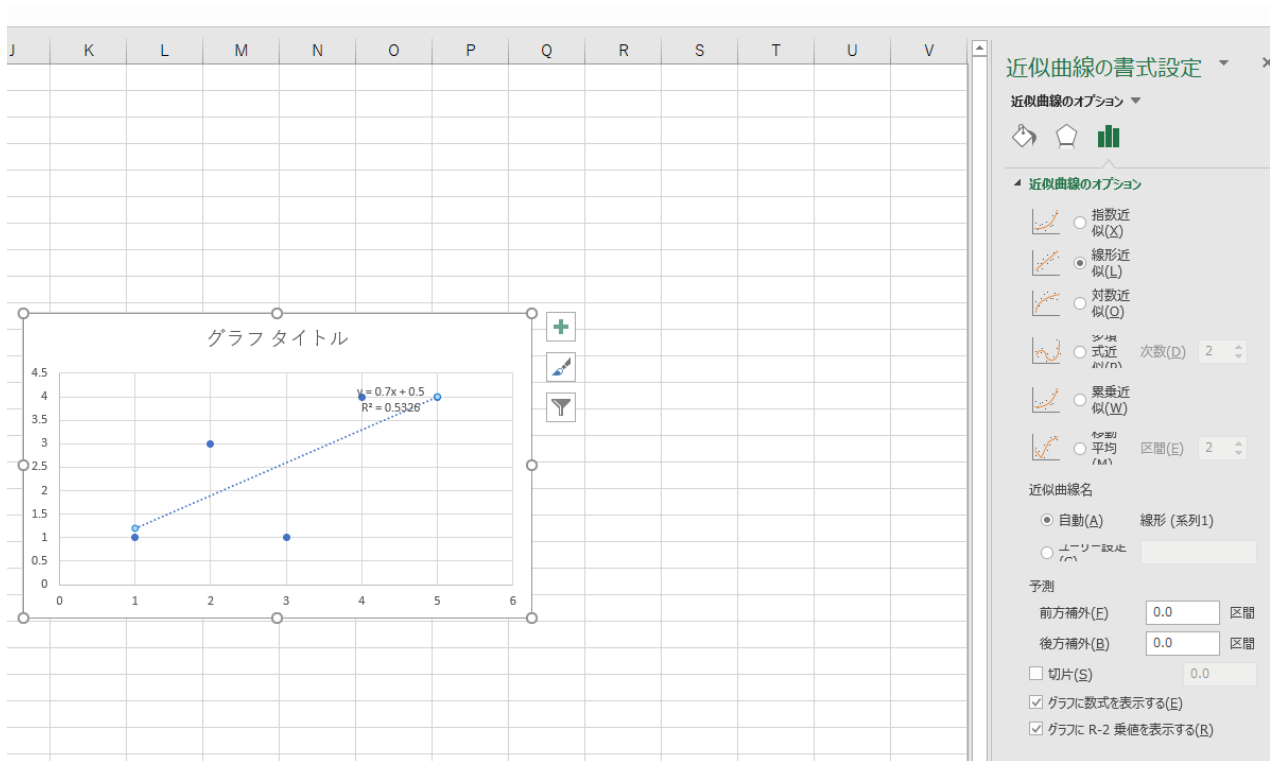


下から2つ目の「近似曲線の追加(R)」を選択して、下記の画面になる。





直線の方程式をグラフ内に表示させるには、右側の近似曲線の書式設定の一番下の「グラフに数式を表示する (E)」にチェックを入れる。決定係数を表示させるには「グラフに R-2 乗値を表示する (R)」にチェックを入れる。下の画面になる。

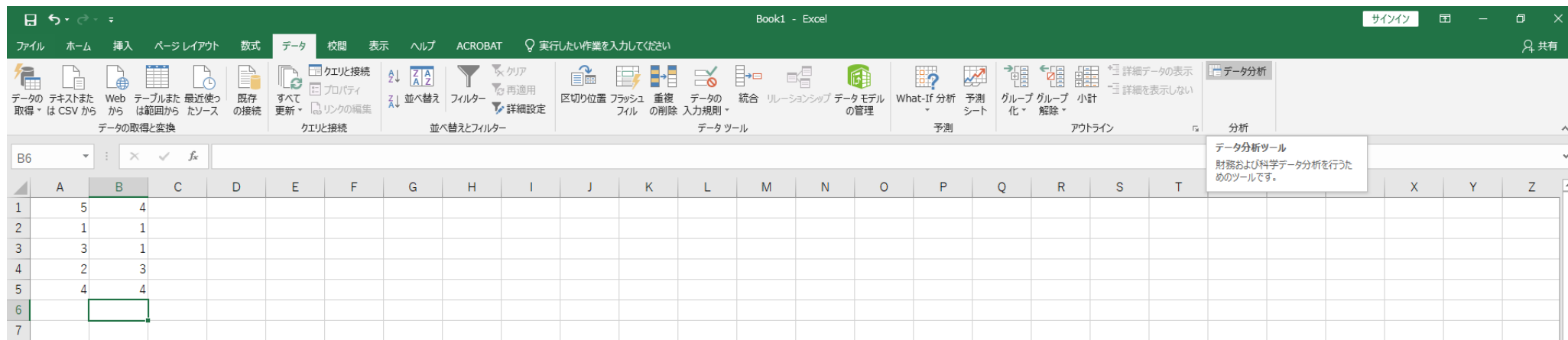


直線の式は  $Y=0.7X+0.5$ , 決定係数  $R^2=0.5326$  と追加される。

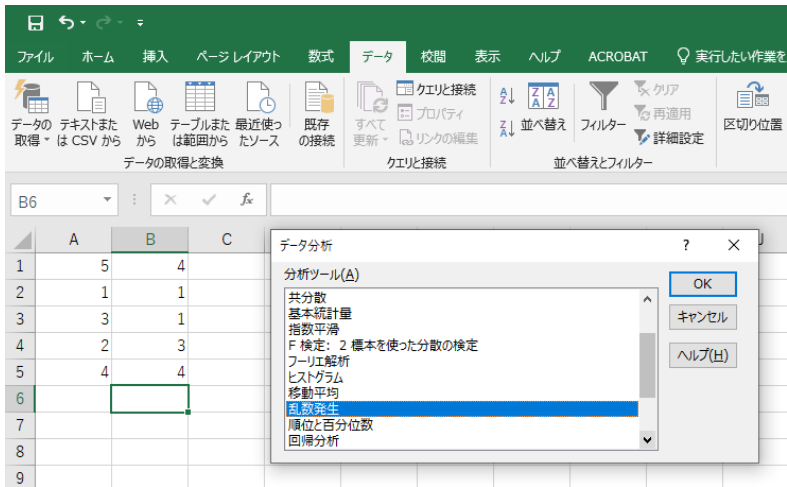
## 3.4.2 「分析ツール」による回帰分析

散布図による方法は、単回帰の場合には、比較的簡単に計算できるが、説明変数が2つ以上の重回帰には適用することは出来なくなる。この場合、「分析ツール」を使うと、簡単に、回帰分析を行うことができる。

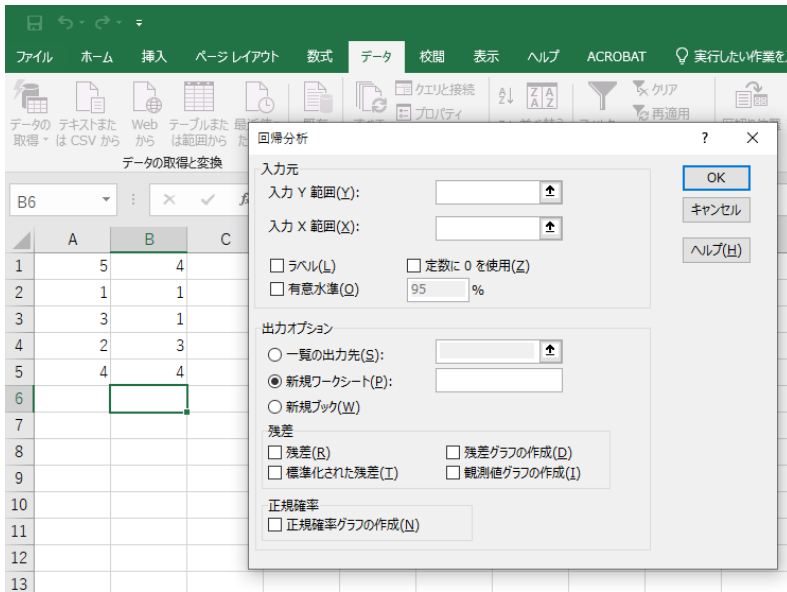
まず、「データ」タブを選ぶ。



「データ分析」のタブをマウスで選択すると、下記のような画面になり、様々なツールが利用できるようになる。主に利用するツールは、「ヒストグラム」と「回帰分析」である。

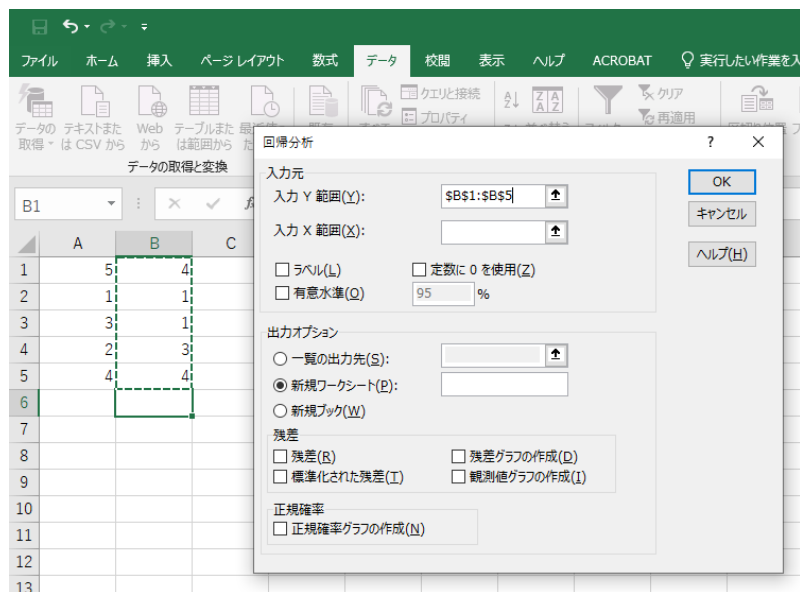


本節では、回帰分析の方法を解説する。まずは、「回帰分析」を選ぶと、下記の画面となる。

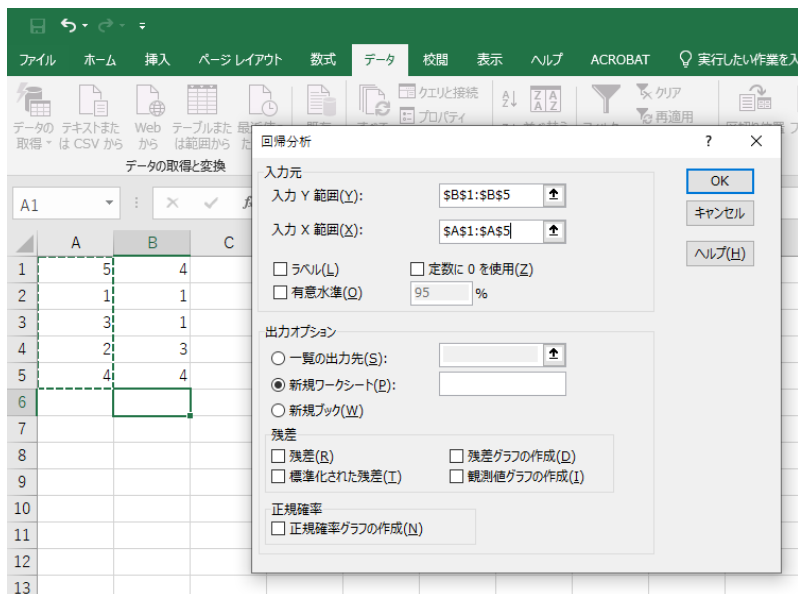


「入力 Y 範囲 (Y)」に B 列のデータ（被説明変数）を選択する。

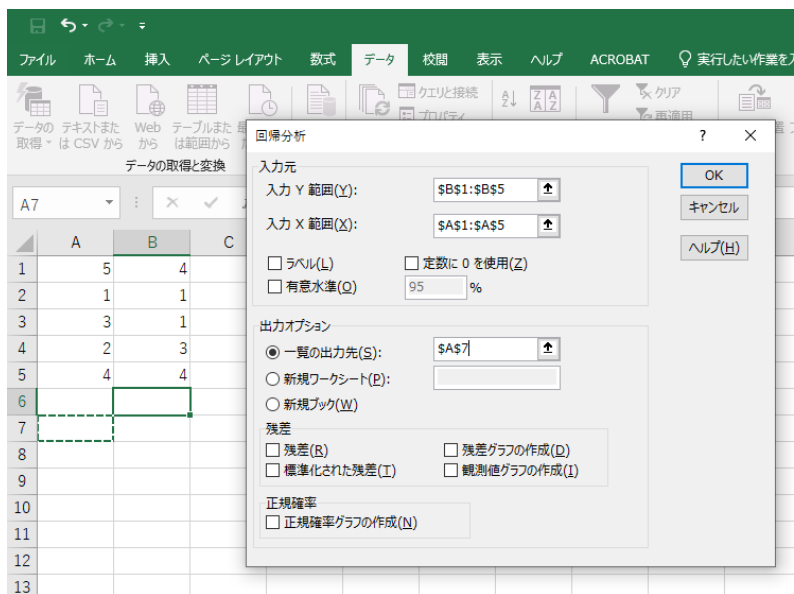
「入力 Y 範囲 (Y)」の右側の空欄をマウスの左ボタンをクリックして、さらに、B1 をマウスの左ボタンでクリック、さらにマウスの左ボタンを押し続けながら B5 でマウスボタンを離す（または、B1:B5 とタイプする）。下記の画面となる。



同様に、「入力 X 範囲 (X)」の右側の空欄をマウスの左ボタンでクリックして、さらに、A1 を左ボタンでクリック、マウスの左ボタンを押し続けながら A5 でマウスボタンを離す（または、A1:A5 と入力する）。下記の画面となる。



「一覧の出力先 (S)」にチェックを入れて、その右側の空欄をマウスの左ボタンでクリック、適当な場所をマウスでクリックして選択する（ここでは、A7 をクリックする。または、A7 とタイプする）。下のような表示になる。



このように入力した後、右側の「OK」ボタンをクリックする。下のような出力結果が得られる。

	A	B	C	D	E	F	G	H	I	J
1	5	4								
2	1	1								
3	3	1								
4	2	3								
5	4	4								
6										
7	概要									
8										
9	回帰統計									
10	重相関 R	0.7298								
11	重決定 R2	0.532609								
12	補正 R2	0.376812								
13	標準誤差	1.197219								
14	観測数	5								
15										
16	分散分析表									
17		自由度	変動	分散	割られた分	有意 F				
18	回帰	1	4.9	4.9	3.418605	0.161594				
19	残差	3	4.3	1.433333						
20	合計	4	9.2							
21										
22		係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%	
23	切片	0.5	1.255654	0.398199	0.717129	-3.49605	4.496051	-3.49605	4.496051	
24	X 値 1	0.7	0.378594	1.848947	0.161594	-0.50485	1.904855	-0.50485	1.904855	
25										
26										
27										
28										

今までの授業では、下記の水色部分を扱った。



The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J
1		5	4							
2		1	1							
3		3	1							
4		2	3							
5		4	4							
6										
7		概要								
8										
9		回帰統計								
10	重相関 R	0.7298								
11	重決定 R2	0.532609								
12	補正 R2	0.376812								
13	標準誤差	1.197219								
14	観測数	5								
15										
16		分散分析表								
17		自由度	変動	分散	F	観測された F	有意 F			
18	回帰	1	4.9	4.9	3.418605	0.161594				
19	残差	3	4.3	1.433333						
20	合計	4	9.2							
21										
22		係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%	
23	切片	0.5	1.255654	0.398199	0.717129	-3.49605	4.496051	-3.49605	4.496051	
24	X 値 1	0.7	0.378594	1.848947	0.161594	-0.50485	1.904855	-0.50485	1.904855	
25										

Excel の「重決定 R2」は決定係数, 「補正 R2」は自由度修正済み決定係数, 「観測数」はデータ数  $n$  のことである。

「残差 + 自由度」の 3, 「合計 + 自由度」の 4 はそれぞれ  $n - k = 5 - 2 = 3$ ,  $n - 1 = 5 - 1 = 4$  であり, 自由度を表す。

また, 「残差 + 変動」の 4.3, 「合計 + 変動」の 9.2 という数字は, それぞれ残差平方和,  $Y$  の平均からの差の

二乗和で、次のものである。

$$\sum_{i=1}^n \hat{u}_i^2 = 4.3 \quad \sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n Y_i^2 - n\bar{Y}^2 = 43 - 5 \times 2.6^2 = 9.2$$

「切片+係数」の0.5, 「X値1+係数」の0.7は, 切片, 傾きを表す ( $Y=0.7X+0.5$ )。

得られた数値と今回得られた数値を比較すると, それぞれの数字がどのような意味かがわかるだろう。

### 3.4.3 決定係数 $R^2$ について

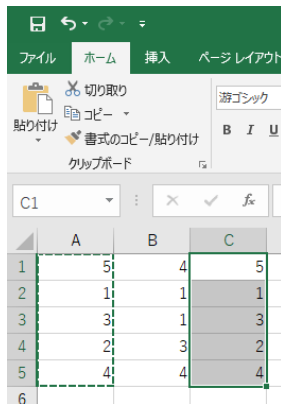
● 説明変数を増やせば、必ず決定係数  $R^2$  は大きくなることを確認する。

都合により、A列のデータ（説明変数）をC列にコピーする。

コピーの方法としては、A1にマウスを持っていき、マウスの左ボタンを押し続けて、A5で左ボタンを離す。

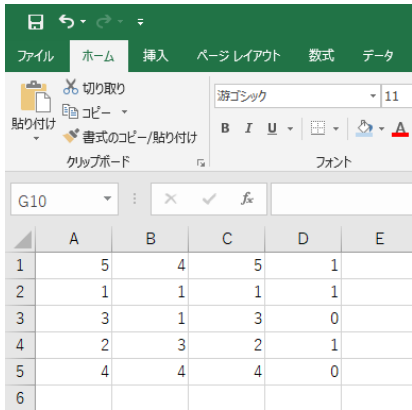
次に、A5にマウスがある状態で、マウスの右ボタンを押し、「コピー (C)」を選択する。C1で右ボタンを押し、「貼り付けのオプション」の一番左のアイコン「貼り付け (P)」を選ぶと、下記のように、A列がC列に

コピーできる。

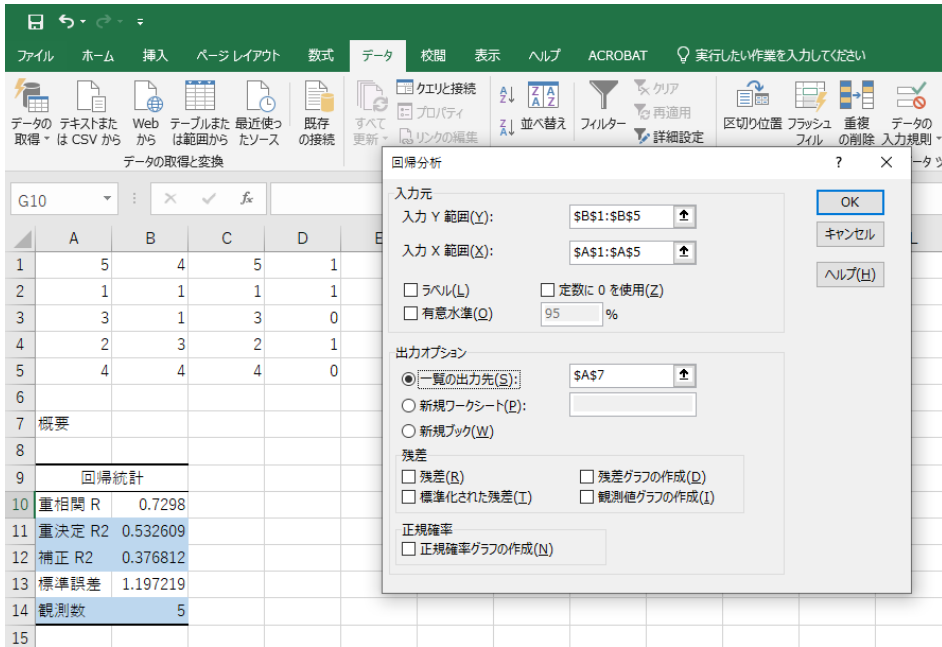


次に、D列に適当に、例えば、1, 1, 0, 1, 0というデータを入力する。

B列を被説明変数、C列・D列を説明変数として回帰分析する。



「データ」タブ、「データ分析」、「回帰分析」、「OK」と順番に選択していくと、下記のように前回のものが残ったままになっている。



「入力 X 範囲 (X)」の欄を削除して、C1 にマウスを置いて、マウスの右ボタンを押し続けて、D5 に移動する

(選択範囲を C1 から D5 とする)。下記の画面になる。

Excelの「回帰分析」ダイアログボックスのスクリーンショット。背景にはデータが入ったワークシートが表示されている。ダイアログボックスの「入力元」セクションには、入力 Y 範囲(Y)が \$B\$1:\$B\$5、入力 X 範囲(X)が \$C\$1:\$D\$5 と設定されている。出力オプションとして「一覧の出力先(S)」が \$A\$7 に設定されている。その他のオプションはほとんど未チェック状態である。

	A	B	C	D	E
1	5	4	5	1	
2	1	1	1	1	
3	3	1	3	0	
4	2	3	2	1	
5	4	4	4	0	
6					
7	概要				
8					
9	回帰統計				
10	重相関 R	0.7298			
11	重決定 R2	0.532609			
12	補正 R2	0.376812			
13	標準誤差	1.197219			
14	観測数	5			
15					

次に、「一覧の出力先(S)」の欄を削除して、例えば、A26 でマウスの左ボタンを押す。

下記の画面となる。

データの取得と変換

回帰分析

入力元  
入力 Y 範囲 (Y): \$B\$1:\$B\$5  
入力 X 範囲 (X): \$C\$1:\$D\$5  
 ラベル (L)     定数に 0 を使用 (Z)  
 有意水準 (Q): 95 %

出力オプション  
 一覧の出力先 (S): \$A\$2:\$G\$14  
 新規ワークシート (P):  
 新規ブック (W)

残差  
 残差 (R)     残差グラフの作成 (D)  
 標準化された残差 (I)     観測値グラフの作成 (I)

正規確率  
 正規確率グラフの作成 (N)

	自由度	変動	分散	割られた分散	有意 F
回帰	1	4.9	4.9	3.418605	0.161594
残差	3	4.3	1.433333		
合計	4	9.2			

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	0.5	1.255654	0.398199	0.717129	-3.49605	4.496051	-3.49605	4.496051
X 値 1	0.7	0.378594	1.848947	0.161594	-0.50485	1.904855	-0.50485	1.904855

右の「OK」ボタンを押す。

A26 以下に下記の結果が出力される。

概要									
	A	B	C	D	E	F	G	H	I
25									
26	概要								
27									
28	回帰統計								
29	重相関 R	0.782718							
30	重決定 R2	0.612648							
31	補正 R2	0.225296							
32	標準誤差	1.334848							
33	観測数	5							
34									
35	分散分析表								
36		自由度	変動	分散	F 値	有意 F			
37	回帰	2	5.636364	2.818182	1.581633	0.387352			
38	残差	2	3.563636	1.781818					
39	合計	4	9.2						
40									
41		係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
42	切片	-0.23636	1.808885	-0.13067	0.907996	-8.01937	7.546642	-8.01937	7.546642
43	X 値 1	0.781818	0.440886	1.77329	0.218182	-1.11516	2.678796	-1.11516	2.678796
44	X 値 2	0.818182	1.272727	0.642857	0.58618	-4.65792	6.294285	-4.65792	6.294285
45									
46									
47									
48									

D 列の変数を Z とすると,

$$Y_i = -0.236 + 0.782 X_i + 0.818 Z_i$$

という結果となった。

D 列の説明変数を加えたことにより, 決定係数は 0.5326 から 0.6126 に増えたが, 自由度修正済み決定係数は 0.3768 から 0.2253 へ低下した。





● 決定係数を比較するためには、被説明変数が同じでなければならない。

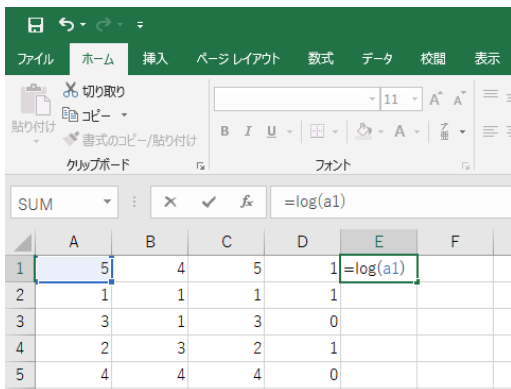
先ほどの例では、

$$Y = 0.5 + 0.7 X \quad R^2 = 0.5326$$

であった。

Y, X に対数を取って、 $\log Y = \alpha + \beta \log X$  を推定してみる。

E 列・F 列に A 列・B 列の対数を求める。E1 に「=log(a1)」とタイプする。



	A	B	C	D	E	F
1	5	4	5	1	=log(a1)	
2	1	1	1	1		
3	3	1	3	0		
4	2	3	2	1		
5	4	4	4	0		

Enter キーを押す。

	A	B	C	D	E	F
1	5	4	5	1	0.69897	
2	1	1	1	1		
3	3	1	3	0		
4	2	3	2	1		
5	4	4	4	0		

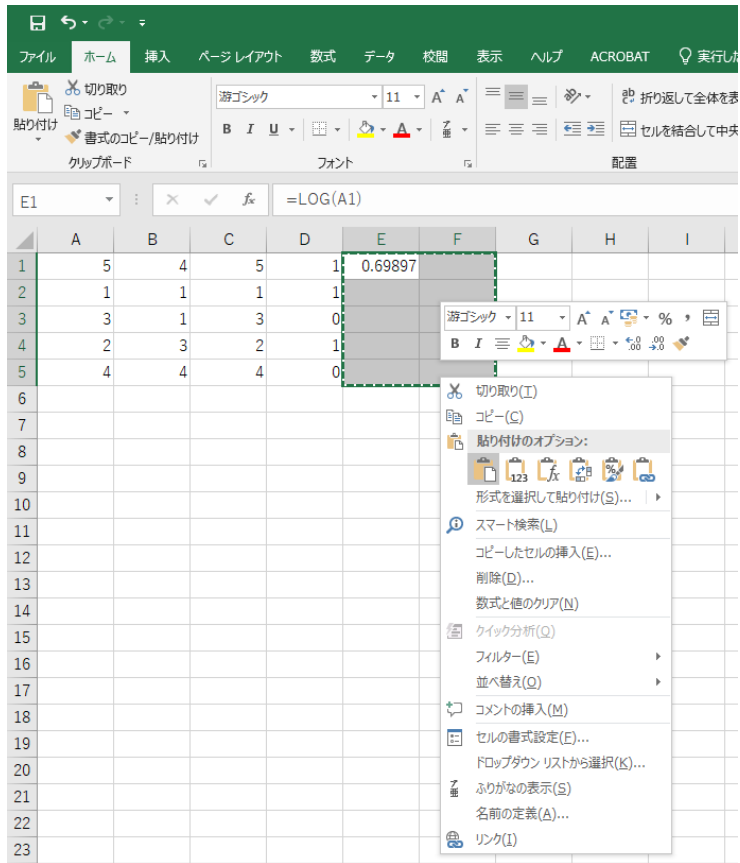
5 の常用対数の値（底が 10，すなわち， $\log_{10} 5$ ）が E1 に計算される。

E1 にマウスを置いて，マウスの右ボタンを押して，「コピー(C)」を選択する。

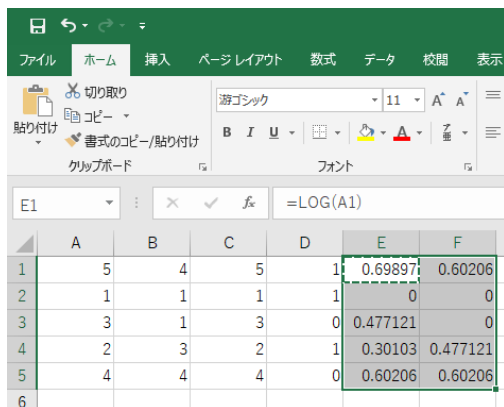
マウスを押し続けながら，F5 で，マウスの右ボタンを離すと，下記のようにになる。

	A	B	C	D	E	F
1	5	4	5	1	0.69897	
2	1	1	1	1		
3	3	1	3	0		
4	2	3	2	1		
5	4	4	4	0		
6						

すぐに，再度，右ボタンを押すと，下記のようにになる。



「貼り付けオプション:」の一番左を選択すると、下記のように対数が計算される。



「入力 Y 範囲 (Y)」を F1 から F5, 「入力 X 範囲 (X)」を E1 から E5, 「一覧の出力先 (S)」は適当なところ (ここでは, A46) を選択して, 「OK」ボタンを押すと, 下記の結果が得られる。

分散分析表					
	自由度	変動	分散	F	有意 F
回帰	1	0.170315	0.170315	2.35494	0.222445
残差	3	0.216968	0.072323		
合計	4	0.387283			

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	0.025354	0.235602	0.107614	0.921095	-0.72444	0.775143	-0.72444	0.775143
X 値 1	0.747636	0.487192	1.534581	0.222445	-0.80283	2.2981	-0.80283	2.2981

$$\log Y = 0.0254 + 0.7476 \log X \quad R^2 = 0.4398$$

となっている。対数を取る前は,

$$Y = 0.5 + 0.7 X \quad R^2 = 0.5326$$

で,  $R^2$  の比較はできない。係数の意味も異なる (この点は後述)。

## 3.4.4 補足

3.4.3 節の冒頭で、「都合により、A列のデータ（説明変数）をC列にコピーする。」と述べた。

そして、C列・D列を説明変数として回帰分析を行った。

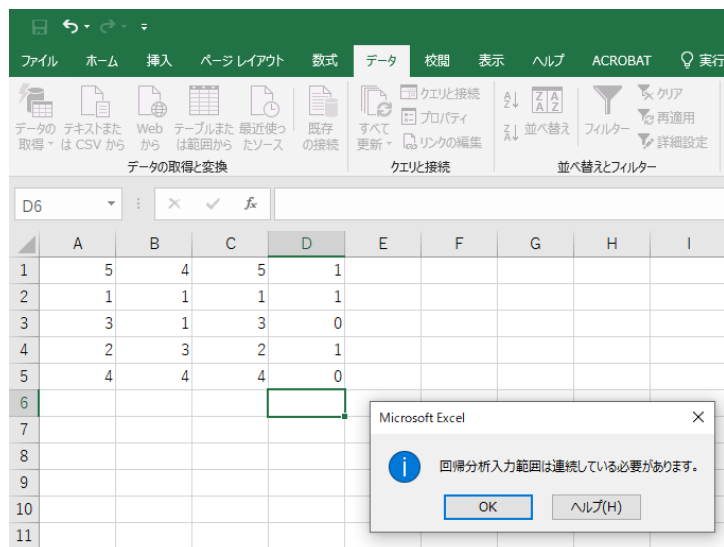
A列とD列を説明変数とするとどうなるかを見る。

「入力 Y 範囲 (Y)」は B 列（これは今までと同様）、「一覧の出力先 (S)」を A7 にする。

「入力 X 範囲 (X)」に、A 列と D 列を選択する（グラフ作成の時と同様に、A1 から A5 までをマウスの左ボタンを押し続けて選択して、次に、Ctrl キーを押しながら D1 から D5 までをマウスの左ボタンを押し続けて選

択する)。

「OK」を押すと、下記の画面になる。



このように、計算結果が出力されない。

「入力 X 範囲(X)」の選択の際には、説明変数データを隣に並べておく必要がある（説明変数が3つであれば、3列連続に並べなければならない）。

これは、試行錯誤で説明変数の種類を変えて、数多くの式を推定する場合はかなり手間がかかる（推定の度に、毎回、説明変数を連続になるように並べ直すことになる）。

この状況を避けるためには、専門の計量経済ソフトを使うことを勧める。

時間の節約にもなり、簡単に推定結果を出すこともできるようになる。

専門の計量経済ソフト：

- ・ 有料 → STATA, EViews, TSP, SPSS など（しかし、高価）
- ・ 無料 → R, Python, Gretl など（ただし、R や Python は若干のプログラミングの知識が必要）

総合的には、Gretl がおすすめ。

<http://gretl.sourceforge.net/>

からダウンロード（windows 版, mac 版あり）

ただし、英語