体験統計学

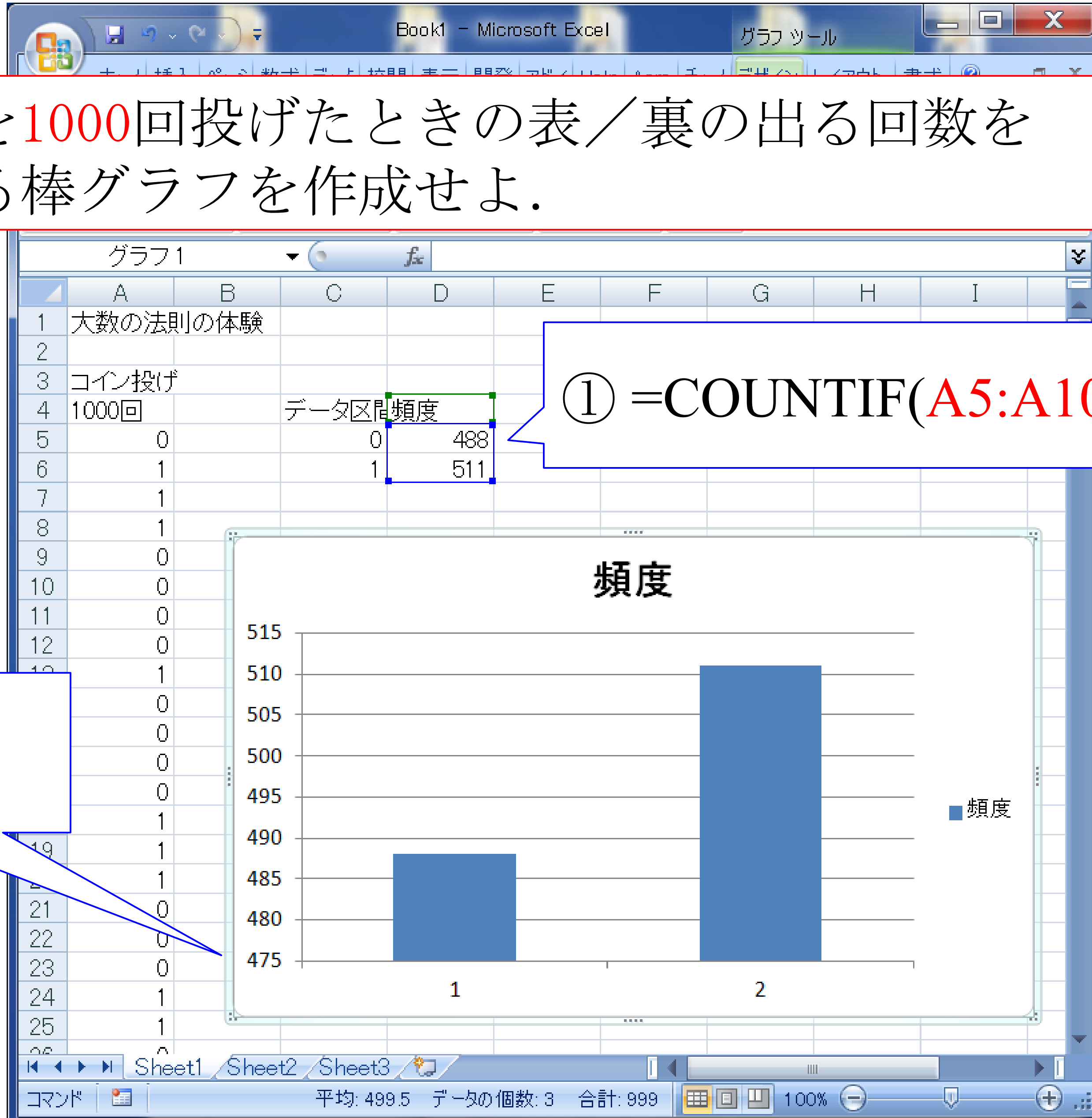
～第4回～

本稿のWebページ

古橋　武

小テスト3-1 解答

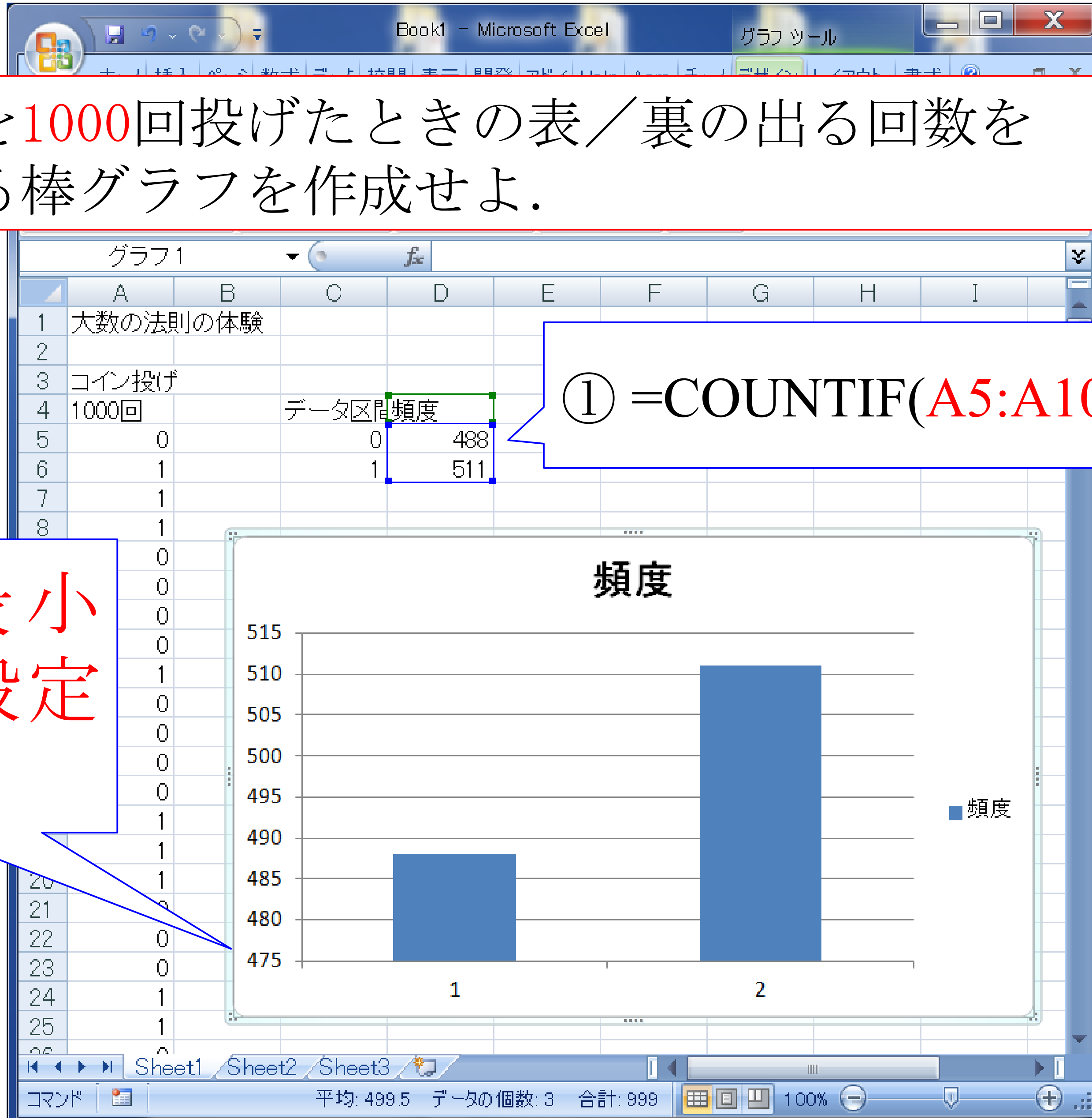
コインを1000回投げたときの表／裏の出る回数を表示する棒グラフを作成せよ.



① =COUNTIF(A5:A1004,0)

小テスト3-1 解答

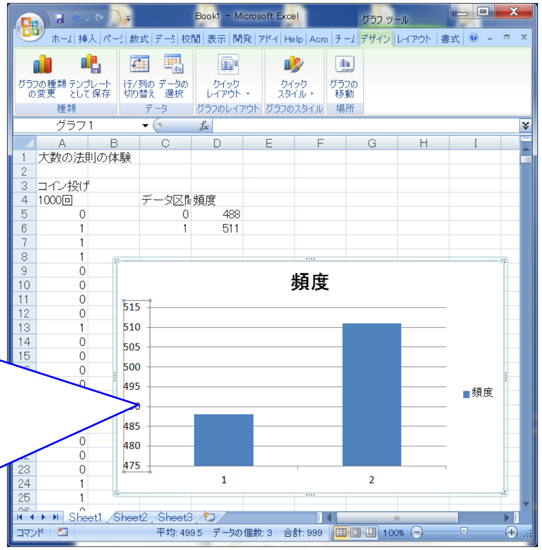
コインを1000回投げたときの表／裏の出る回数を表示する棒グラフを作成せよ.



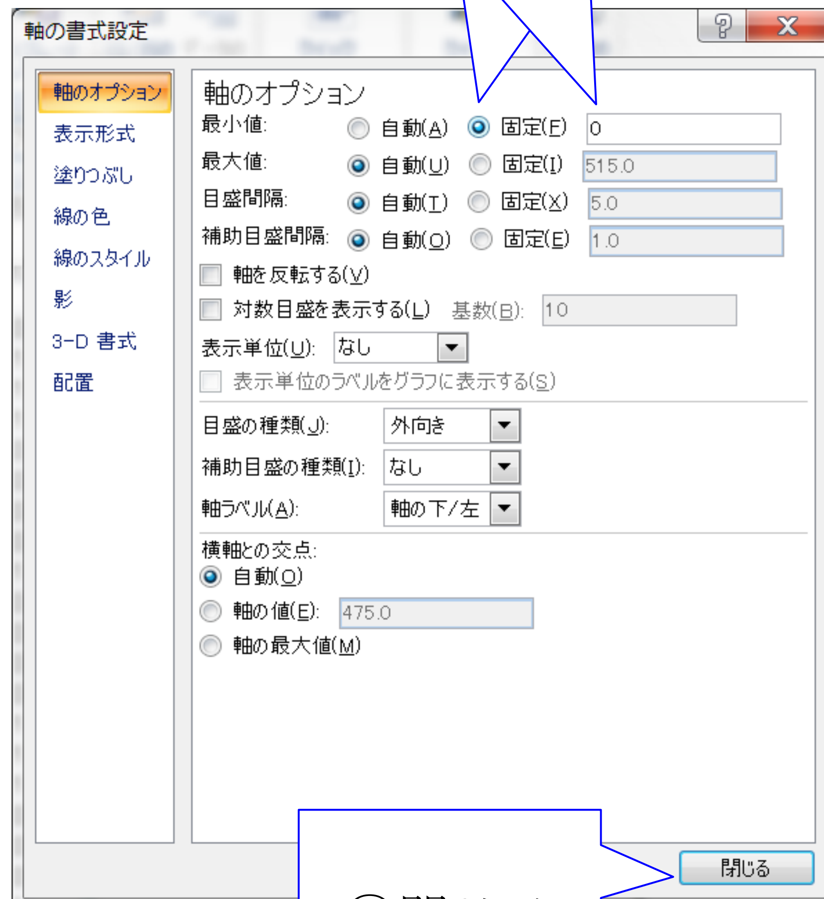
① =COUNTIF(A5:A1004,0)

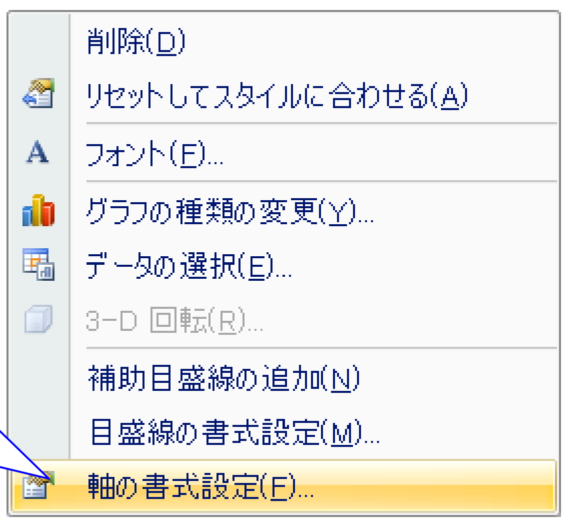
縦軸の最小値を0に設定する

①縦軸の数字の辺りにカーソルを持ってきて左クリック



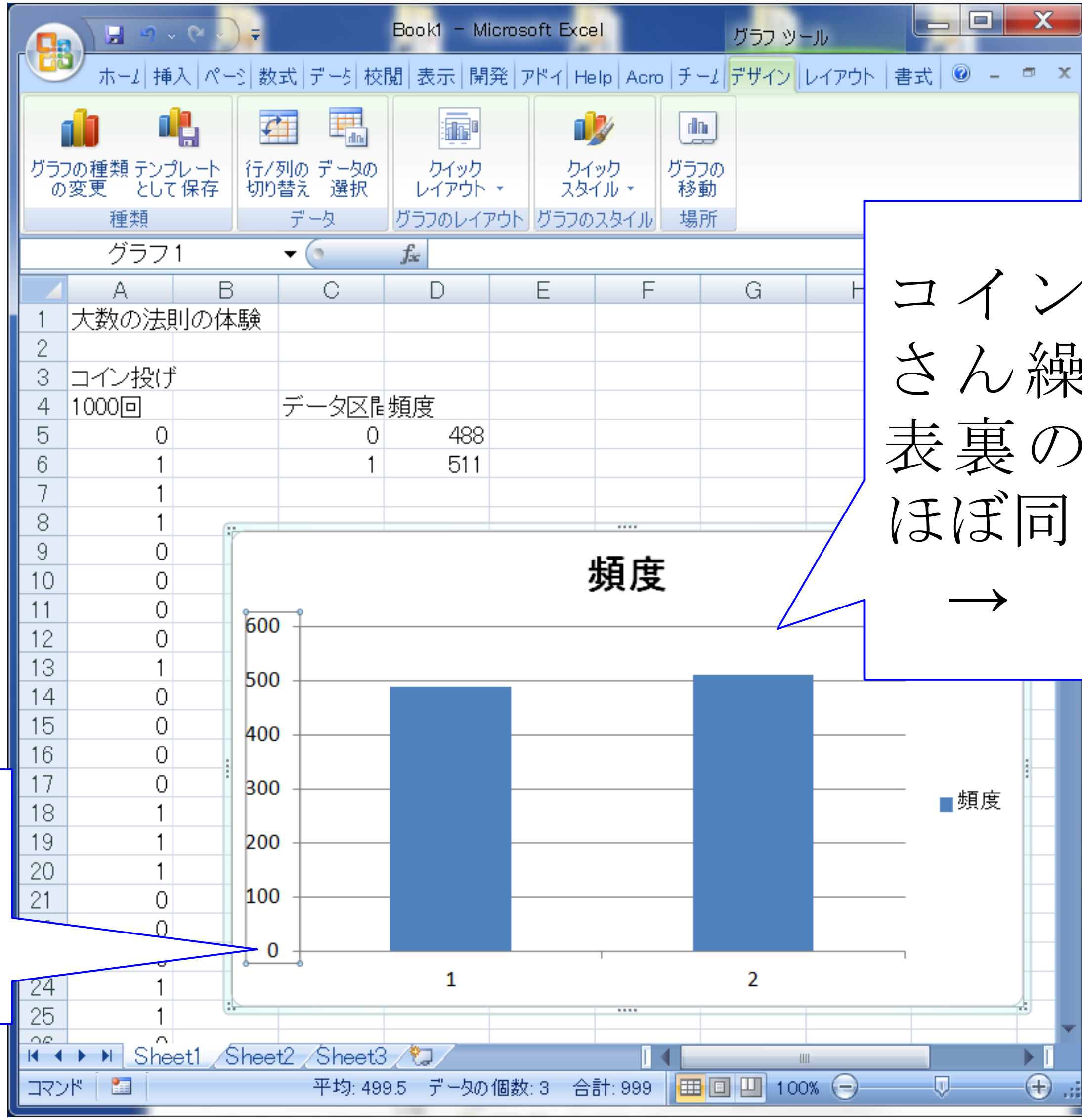
③最小値の固定の選択ボタンをクリックして，0を入力





②軸の書式設定を選択

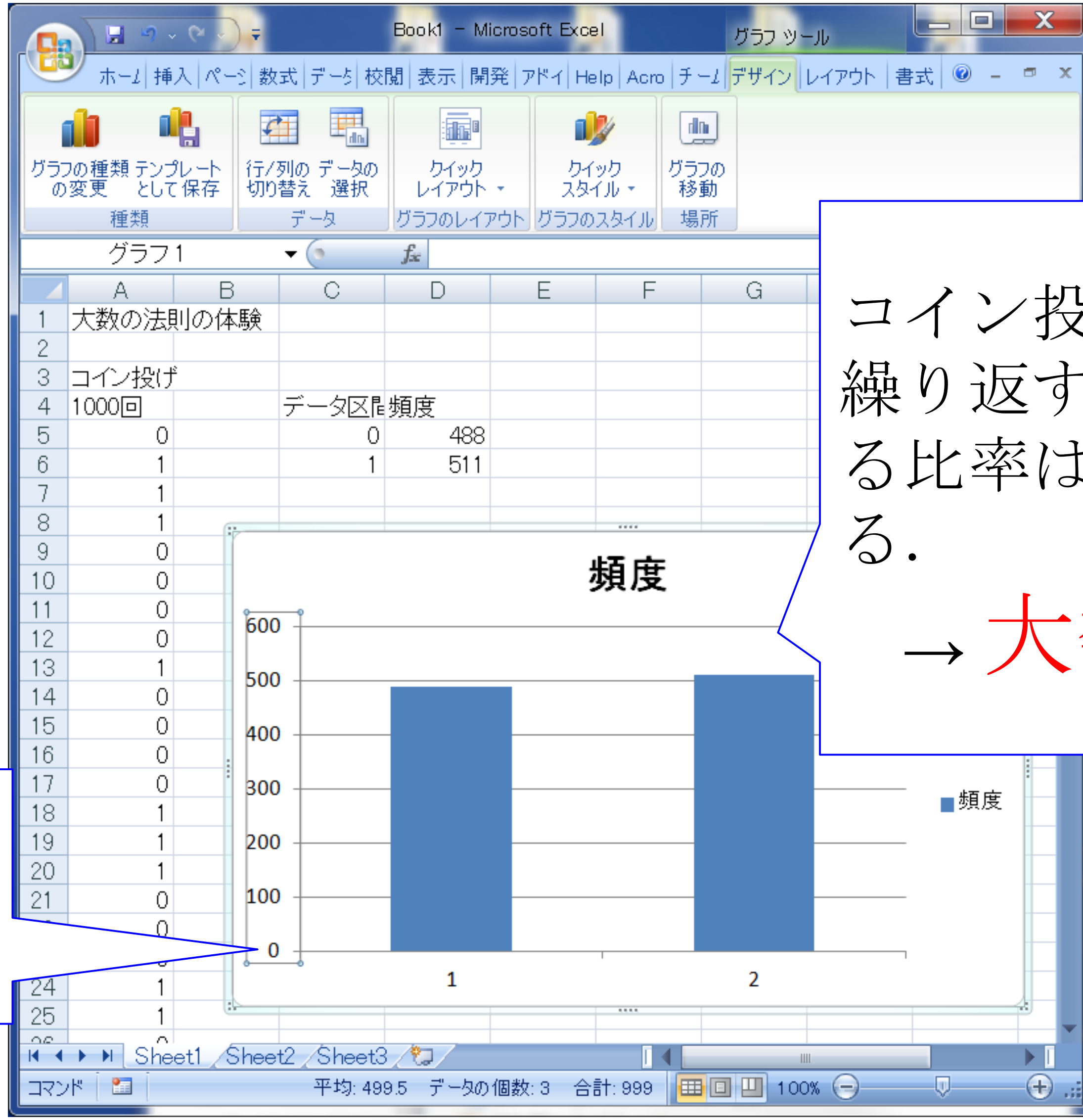
④閉じる



コイン投げをたくさん繰り返すと，表裏の出る比率はほぼ同じとなる．

→

縦軸の最小値を0に設定できた



縦軸の最小値を0に設定できた

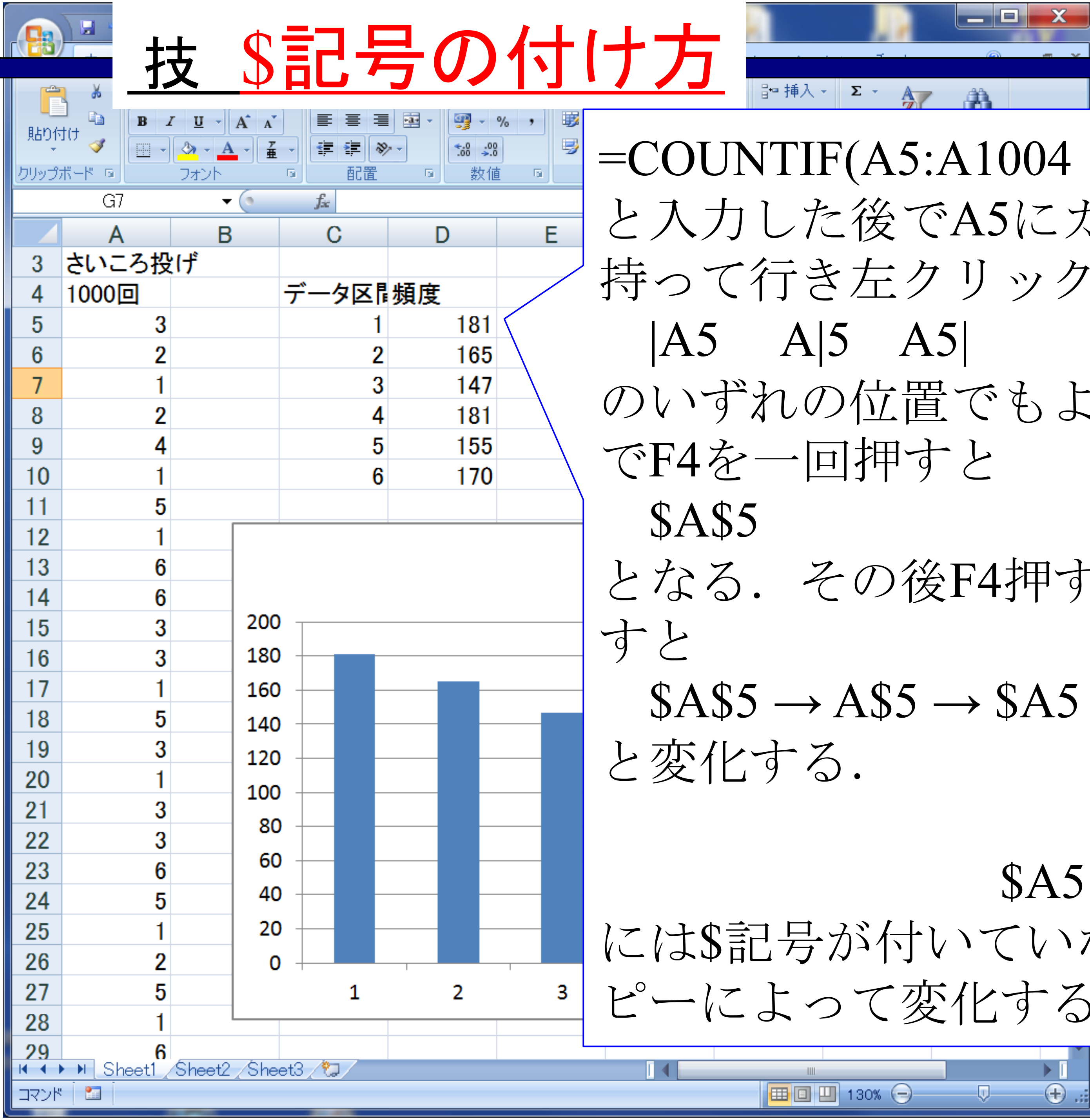
コイン投げをたくさん繰り返すと，表裏の出る比率はほぼ同じとなる．

→ 大数の法則

小テスト3-2　解答

サイコロを1000回投げたときの各目の出る回数を表示する棒グラフを作成せよ.

技　$記号の付け方



=COUNTIF(A5:A1004
と入力した後でA5にカーソルを持って行き左クリック

|A5　A|5　A5|

のいずれの位置でもよい．この状態でF4を一回押すと

$A$5

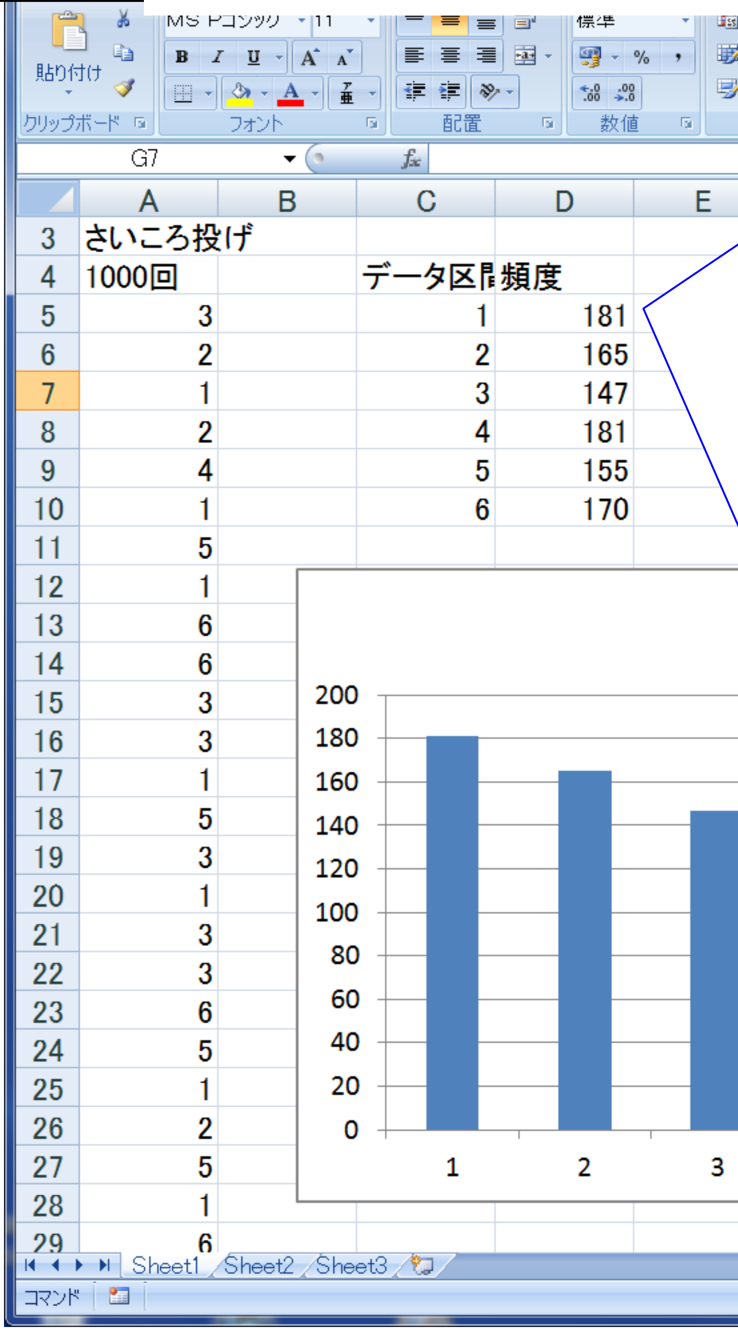
となる．その後F4押すことを繰り返すと

$A$5 → A$5 → $A5 → A5 → $A$5

と変化する．

$A5の場合は，5には$記号が付いていないので，コピーによって変化する．

技　$記号の付け方



=COUNTIF(A5:A1004

とした後でA5にカーソルを持って行き左クリック

|A5　A|5　A5|

のいずれの位置でもよい．この状態でF4を一回押すと

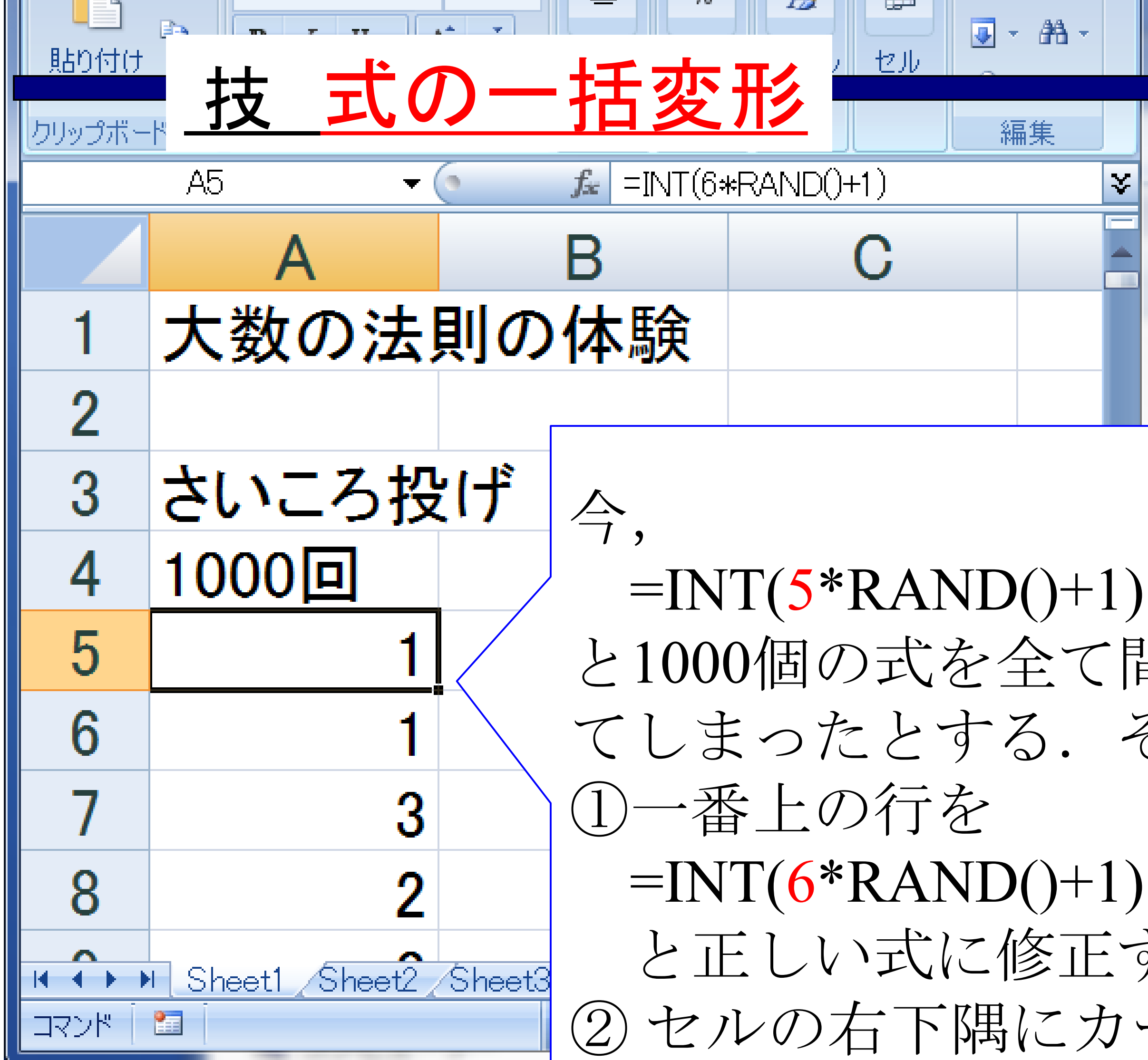
$A$5

となる．その後F4押すことを繰り返すと

$A$5 → A$5 → $A5 → A5 → $A$5

と変化する．$A, $5はそれぞれ$記号が付くことで，コピーによって変化しない．$A5の場合は，5には$記号が付いていないので，コピーによって変化する．

技　式の一括変形



今，

=INT(5*RAND()+1)

と1000個の式を全て間違って入力してしまったとする．そのときは

①一番上の行を

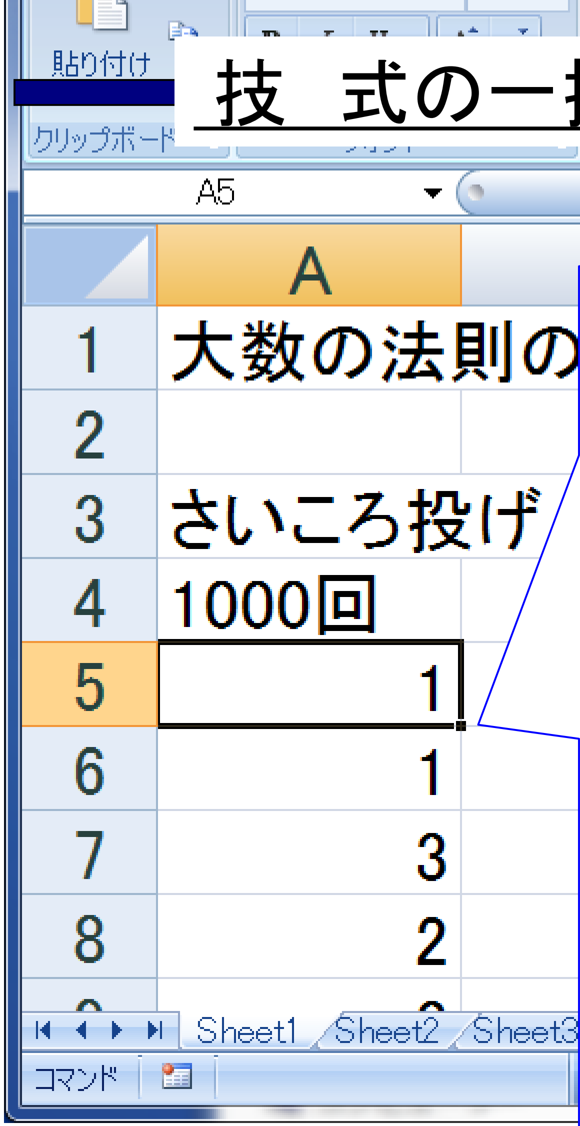
=INT(6*RAND()+1)

と正しい式に修正する．

② セルの右下隅にカーソルを持って行き，カーソルが＋の形に変化したら左ダブルクリック

→

技　式の一括変形



今,

=INT(5*RAND()+1)

と1000個の式を全て間違って入力してしまったとする．そのときは

①一番上の行を

=INT(6*RAND()+1)

と正しい式に修正する．

② セルの右下隅にカーソルを持って行き，カーソルが＋の形に変化したら左ダブルクリック

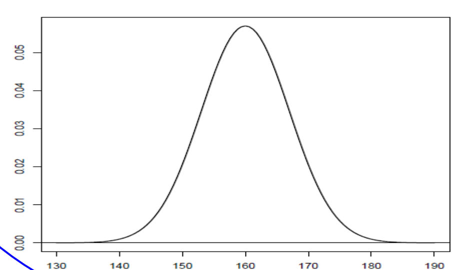
→ これだけで下の999個の式を一括して修正できる．

正規分布

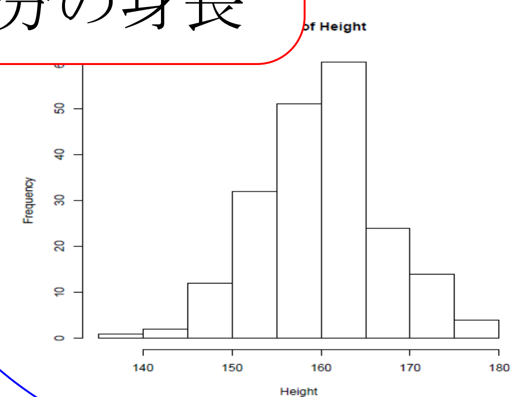
統計学において最も重要な分布は　　　　　　である．多くの理論は母集団の分布が正規分布に従うという仮定の下に成り立っている．また、

高校2年の男子生徒

母集団

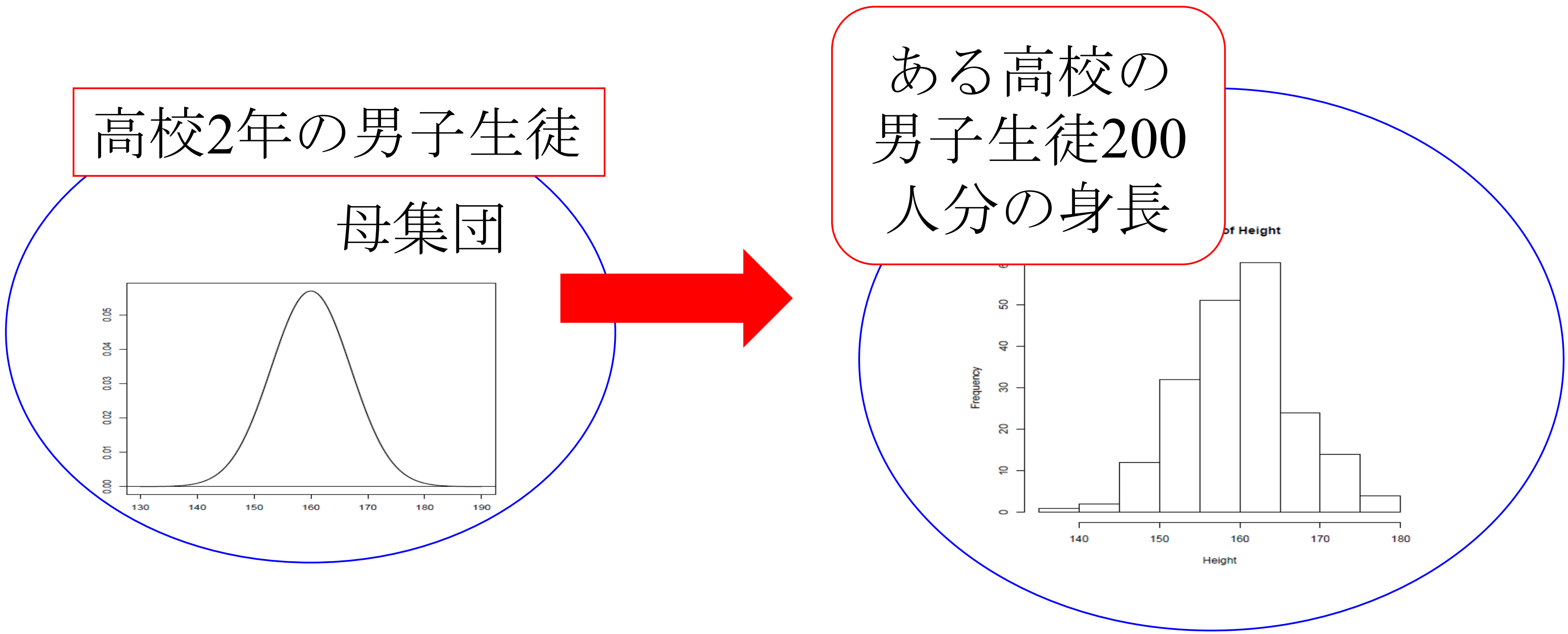


ある高校の男子生徒200人分の身長



正規分布

統計学において最も重要な分布は**正規分布**である．多くの理論は母集団の分布が正規分布に従うという仮定の下に成り立っている．また、

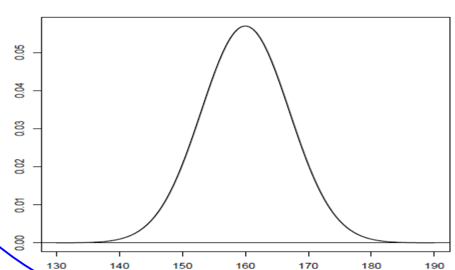


正規分布

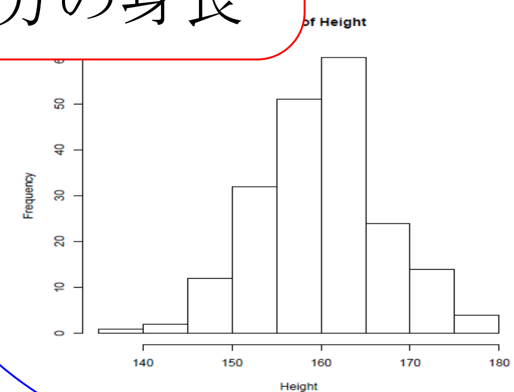
統計学において最も重要な分布は**正規分布**である．多くの理論は母集団の分布が正規分布に従うという仮定の下に成り立っている．また、**この世の多くのデータは正規分布に従うことが経験的に知られている．**

高校2年の男子生徒

母集団



ある高校の男子生徒200人分の身長



中心極限定理

中心極限定理
母集団の分布がどんな分布であっても，
　　　　　は標本の大きさnを増やしたとき近似的に正規分布に従う．　．

コインをn回投げたときに表がq回出たとする．表の出る比率（平均値）をq/nとする．この　を増やすにつれて，この比率は

n個のサンプル

$$a_1, a_1, a_1, \ldots, a_n$$

標本平均

$$\overline{a} = \frac{1}{n}(a_1 + a_2 + \cdots + a_n)$$

中心極限定理

中心極限定理

母集団の分布がどんな分布であっても，**標本平均の分布**は標本の大きさnを増やしたとき近似的に正規分布に従う．

コインをn回投げたときに表がq回出たとする．表の出る比率（平均値）をq/nとする．この　を増やすにつれて，この比率は

n個のサンプル

$$a_1, a_1, a_1, \ldots, a_n$$

標本平均

$$\bar{a} = \frac{1}{n}(a_1 + a_2 + \cdots + a_n)$$

中心極限定理

中心極限定理
母集団の分布がどんな分布であっても，標本平均の分布は標本の大きさnを増やしたとき近似的に正規分布に従う．

コインをn回投げたときに表がq回出たとする．表の出る比率（平均値）をq/nとする．**この*n*を**増やすにつれて，この比率は**正規分布に近づく．**

***n*個のサンプル**

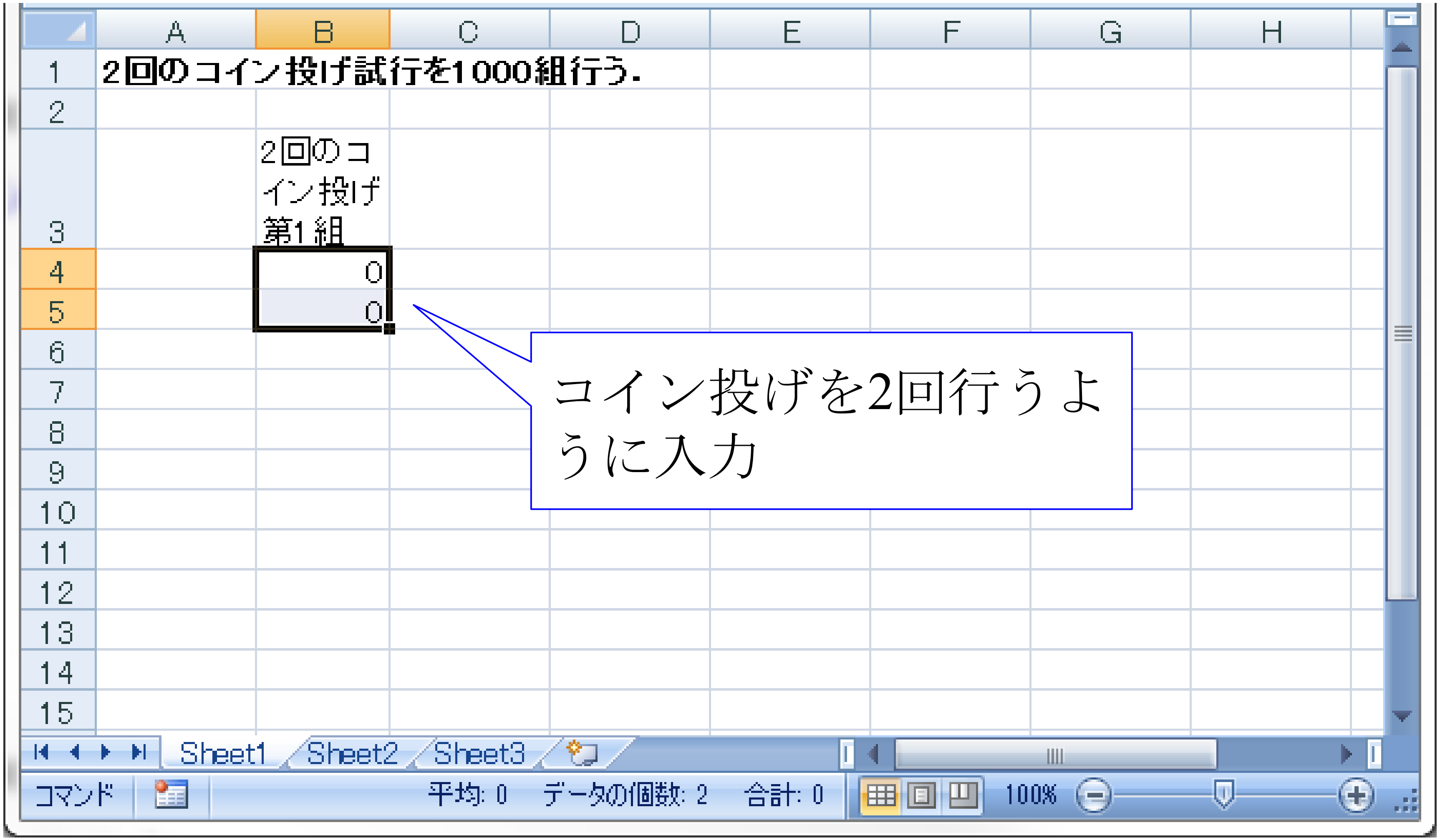
$$a_1, a_1, a_1, \ldots, a_n$$

標本平均

$$\bar{a} = \frac{1}{n}(a_1 + a_2 + \cdots + a_n)$$

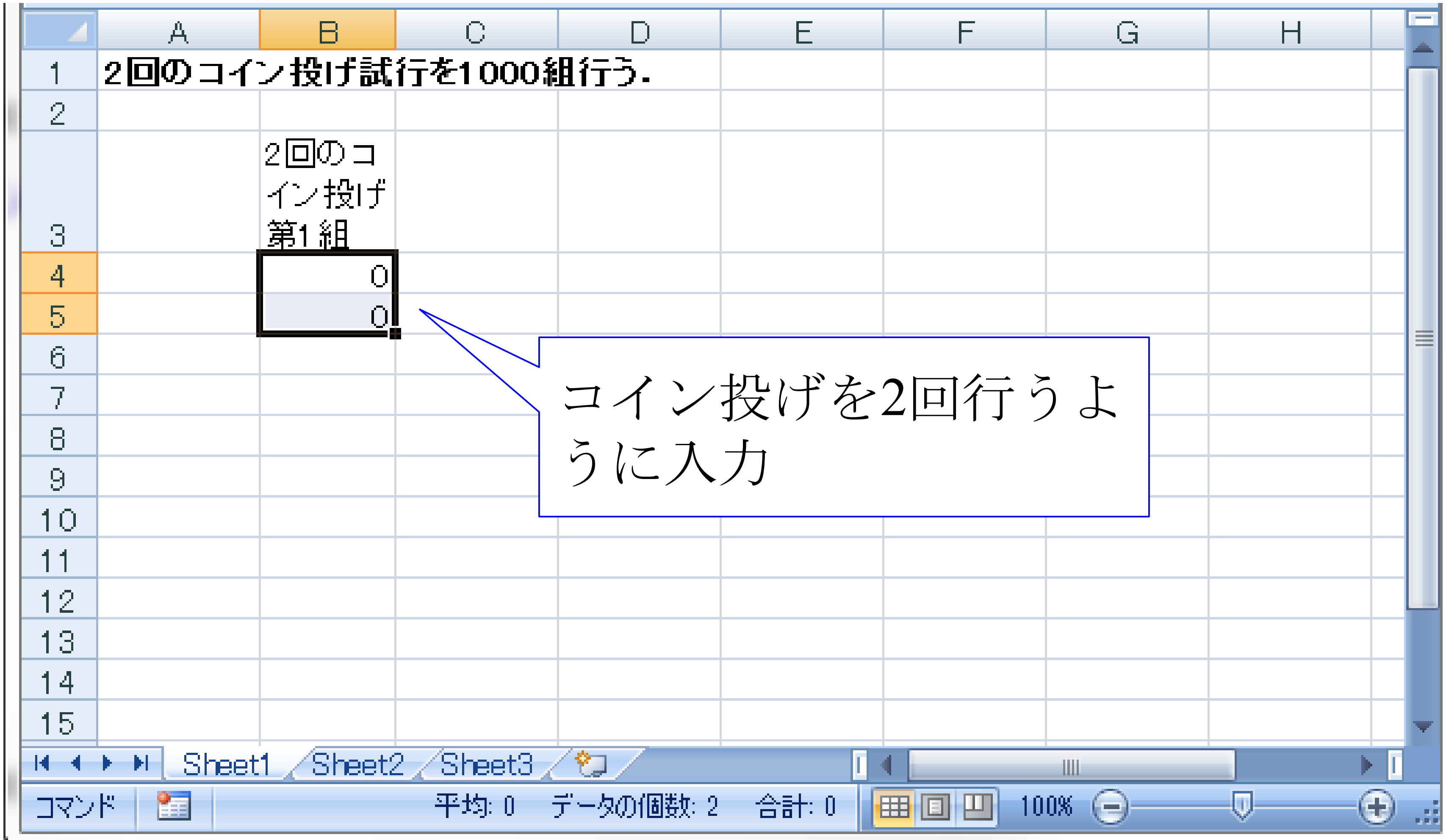
中心極限定理を体験しよう

2回コインを投げたとき（2試行と呼ぶ）に表がq回出たとして，比率$q/2$を求める．2試行を１組とする．1000組の結果について　　　　　　　　　　　　をみよう．

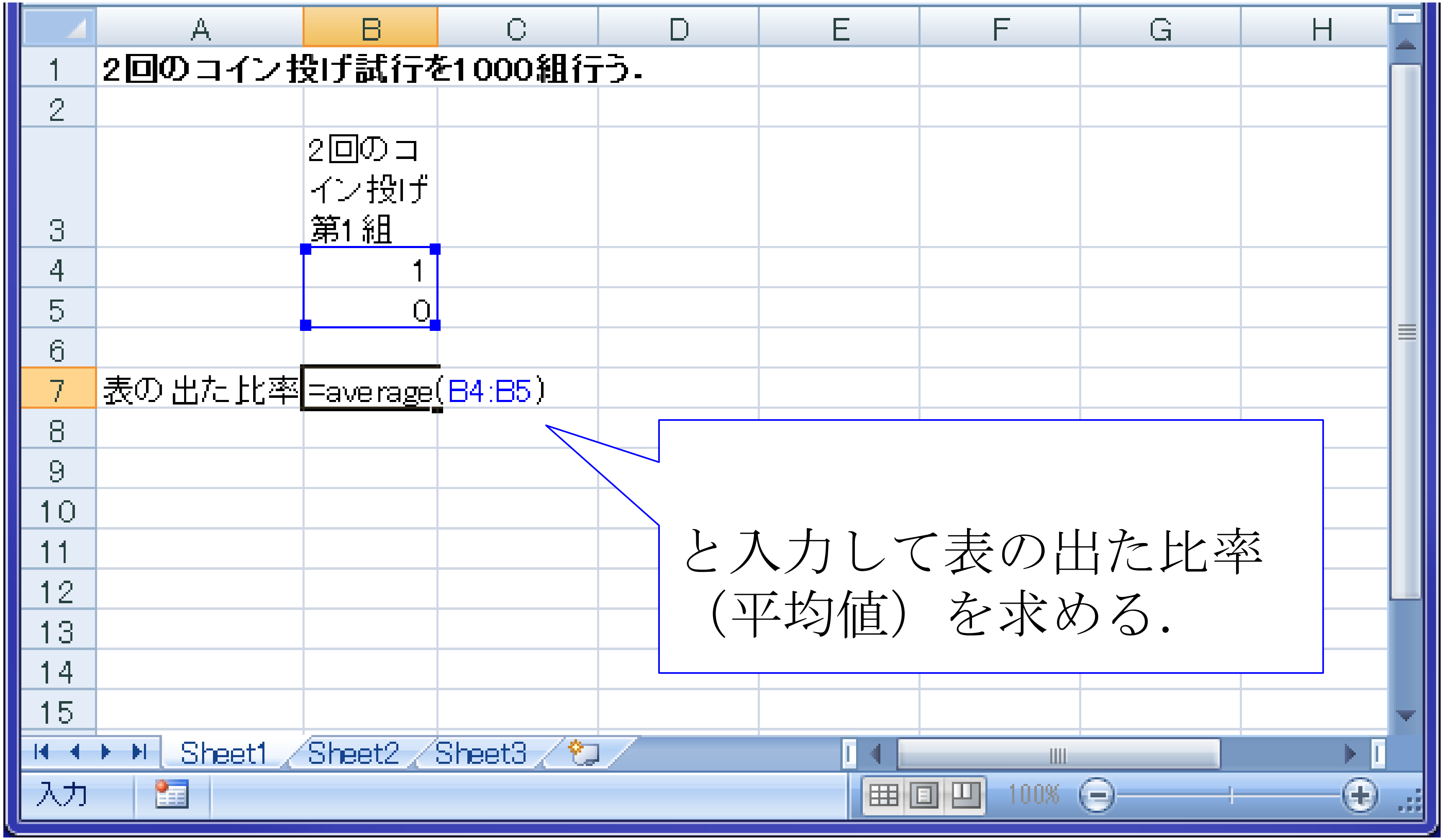


中心極限定理を体験しよう

2回コインを投げたとき（2試行と呼ぶ）に表がq回出たとして，比率$q/2$を求める．2試行を１組とする．1000組の結果について**表の出る比率(平均値)の分布**をみよう．

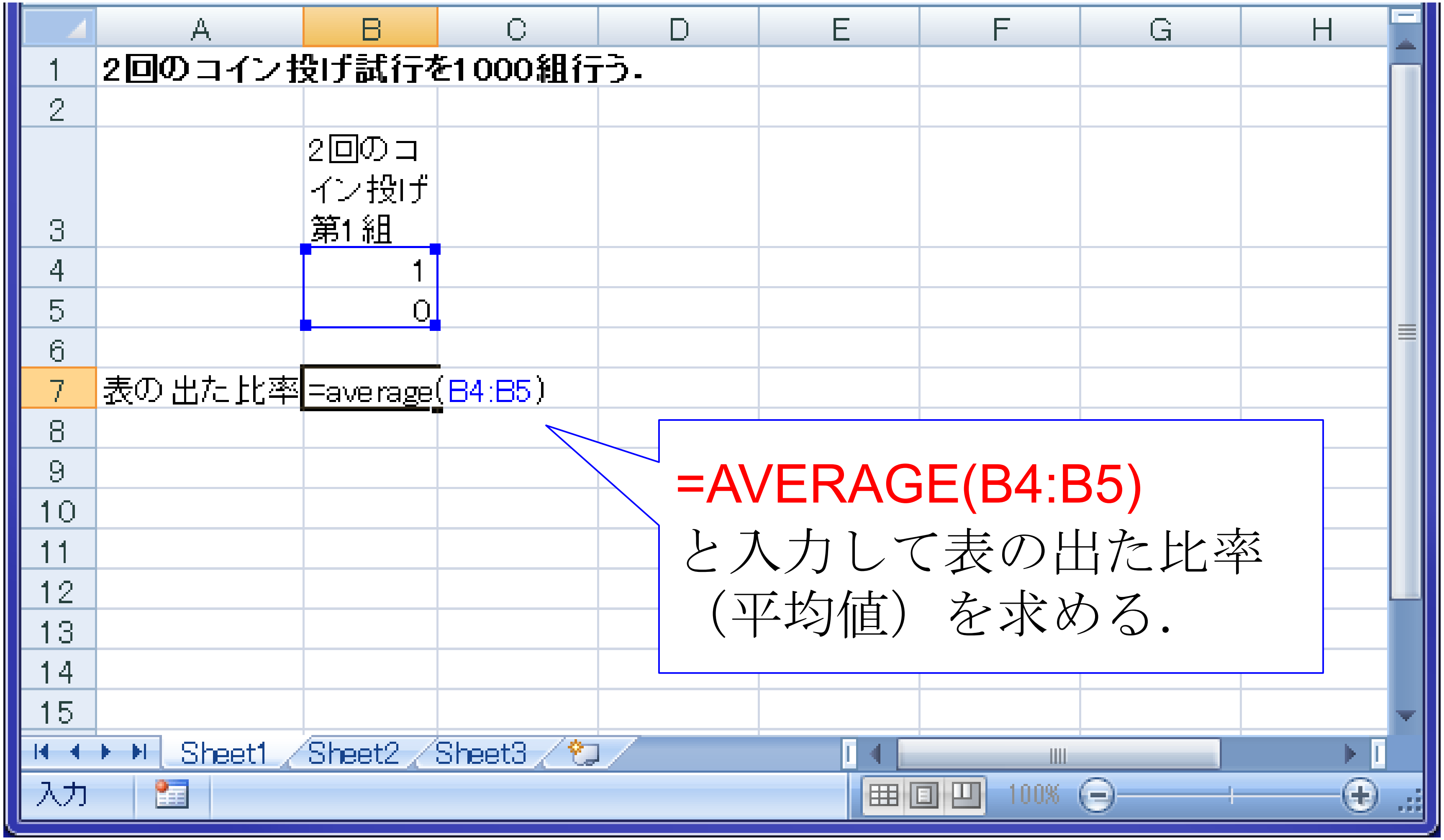


中心極限定理を体験しよう

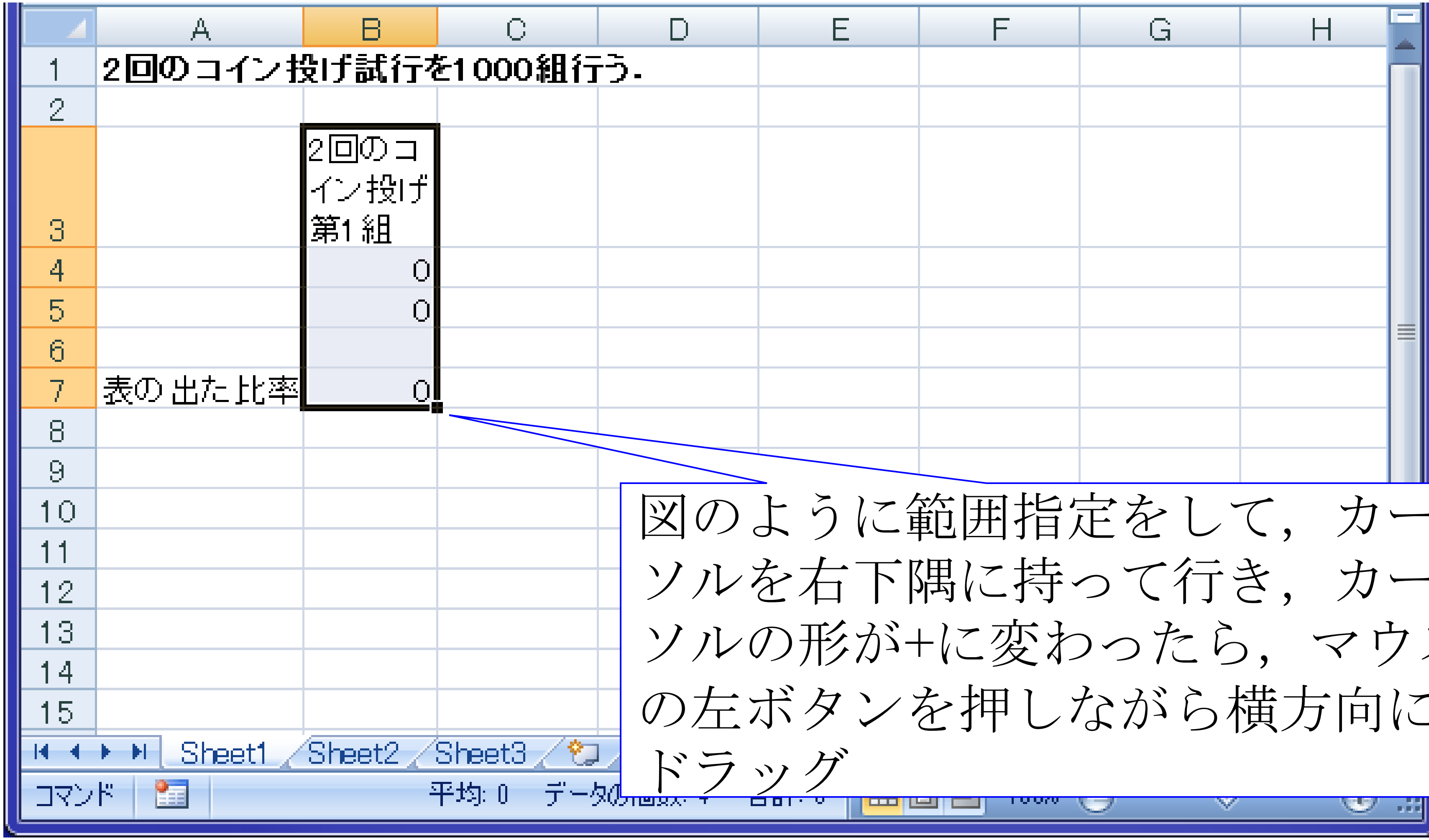


と入力して表の出た比率（平均値）を求める.

中心極限定理を体験しよう

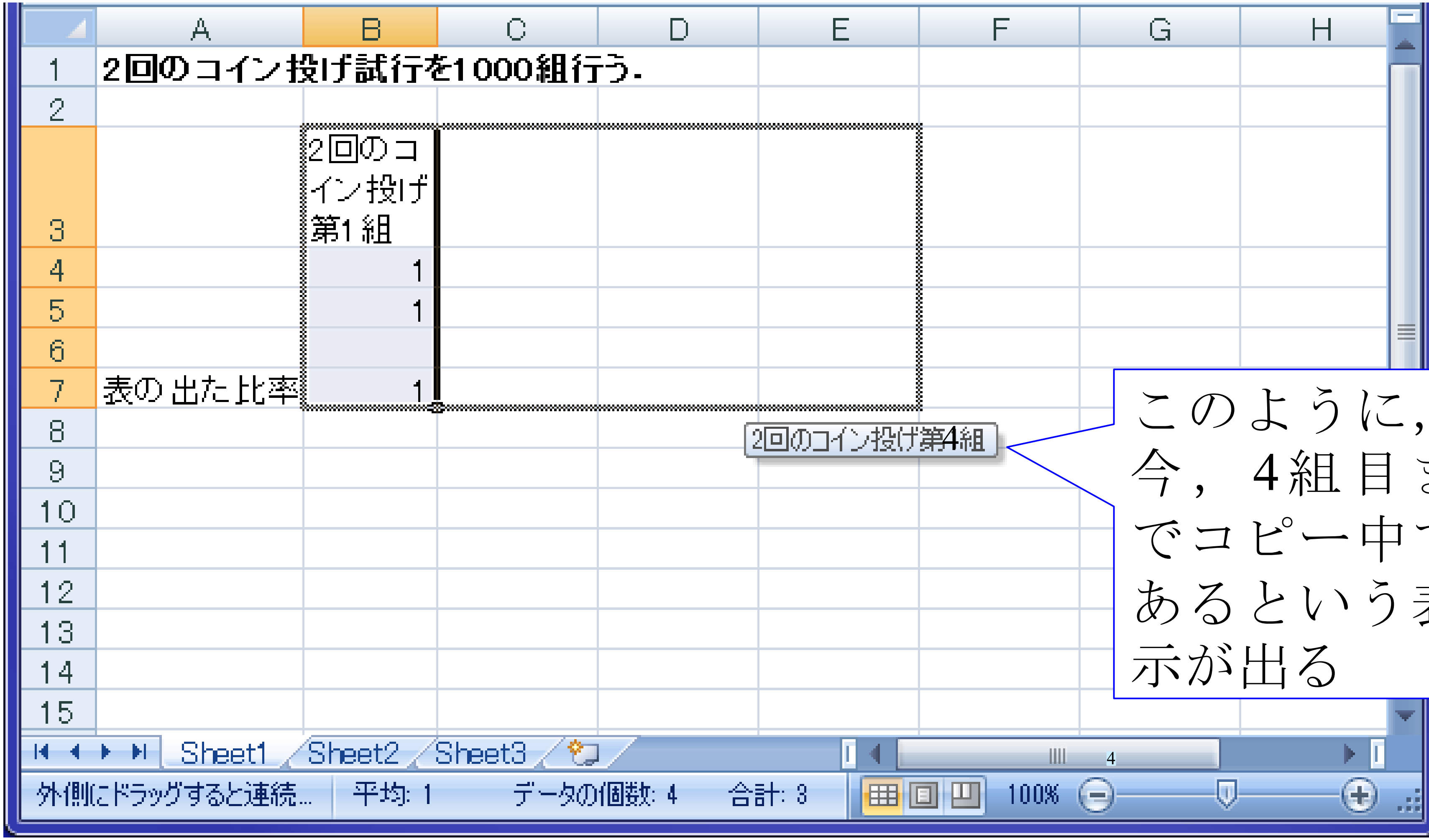


中心極限定理を体験しよう



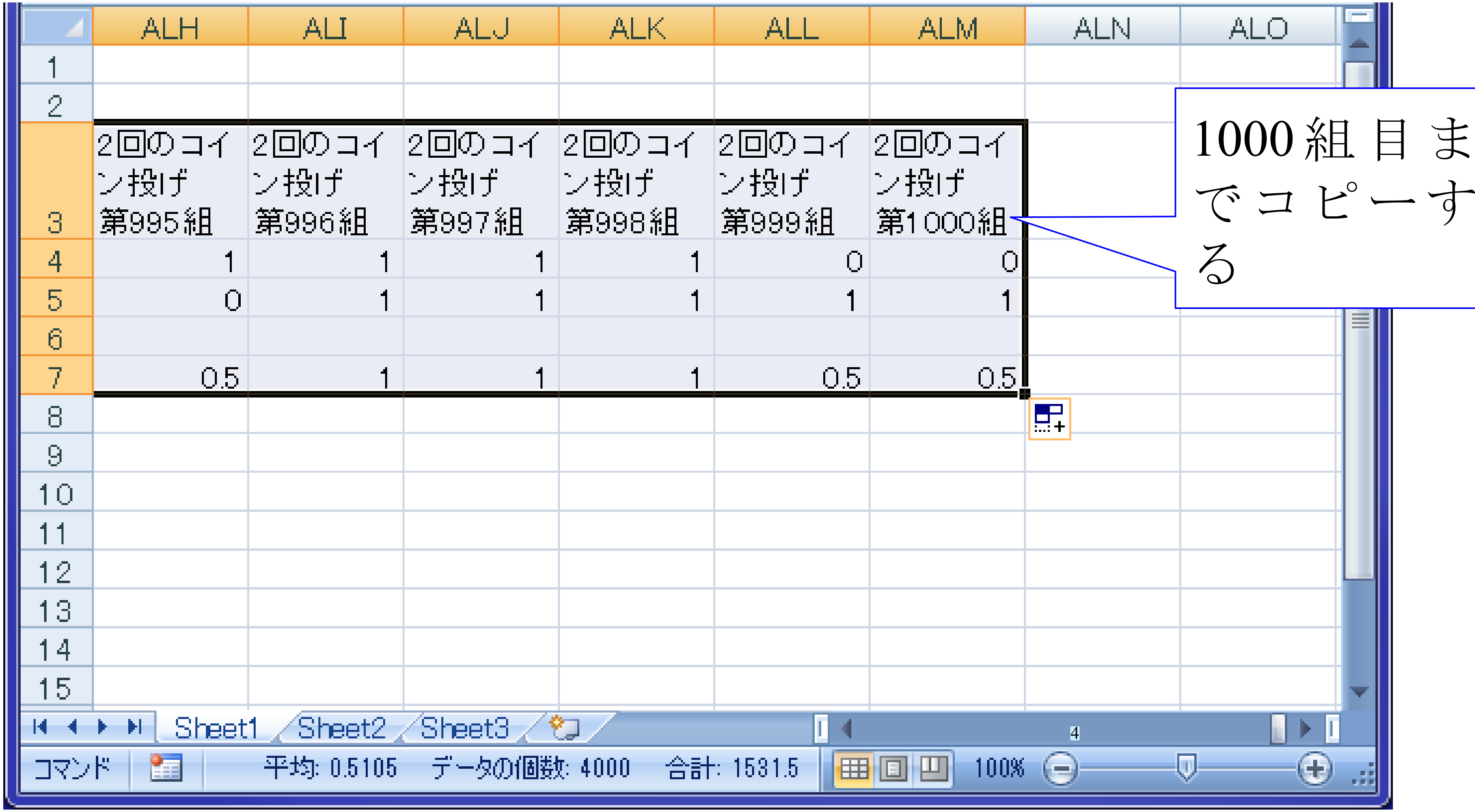
図のように範囲指定をして，カーソルを右下隅に持って行き，カーソルの形が+に変わったら，マウスの左ボタンを押しながら横方向にドラッグ

中心極限定理を体験しよう

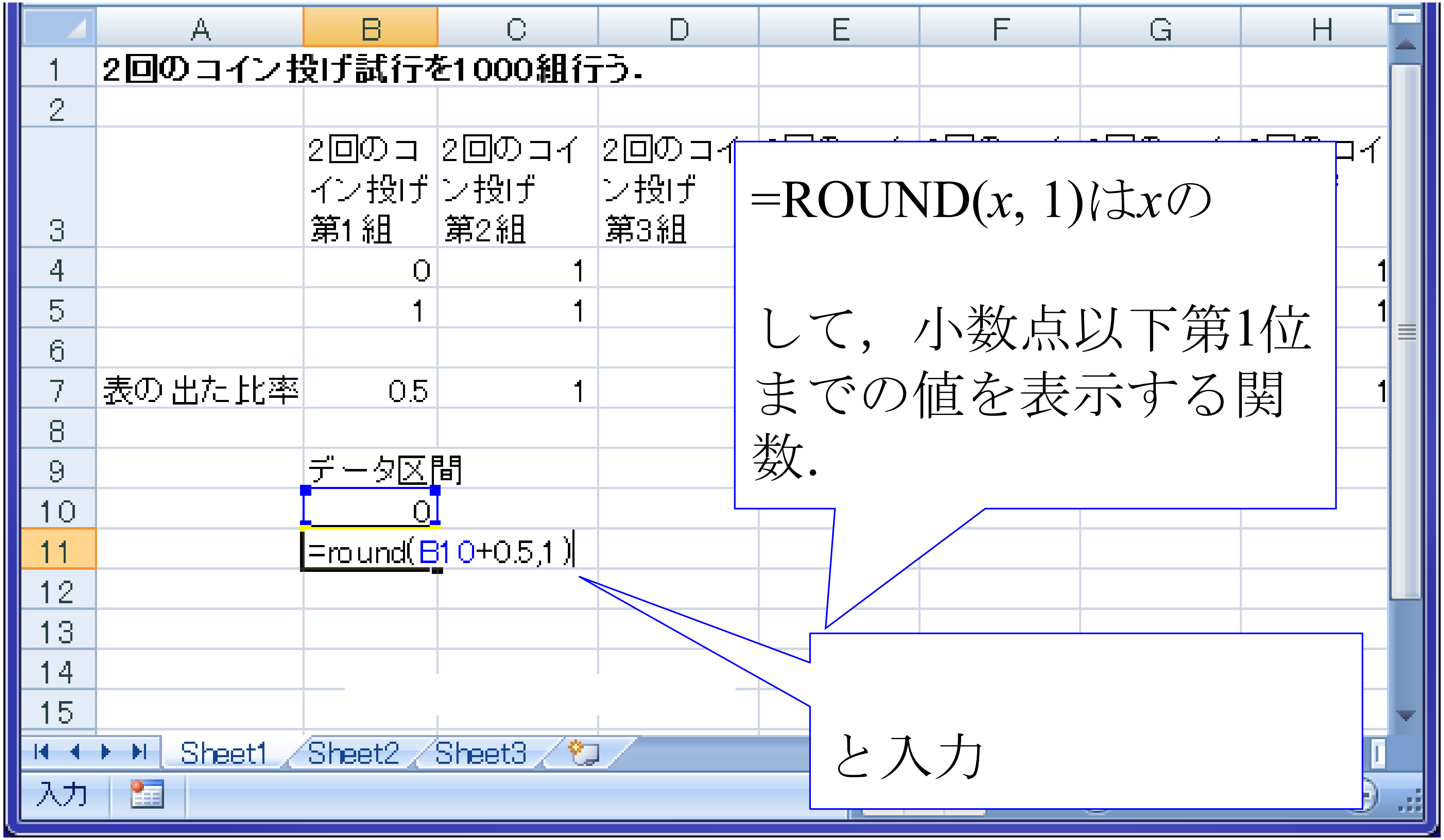


このように，今，4組目までコピー中であるという表示が出る

中心極限定理を体験しよう



中心極限定理を体験しよう

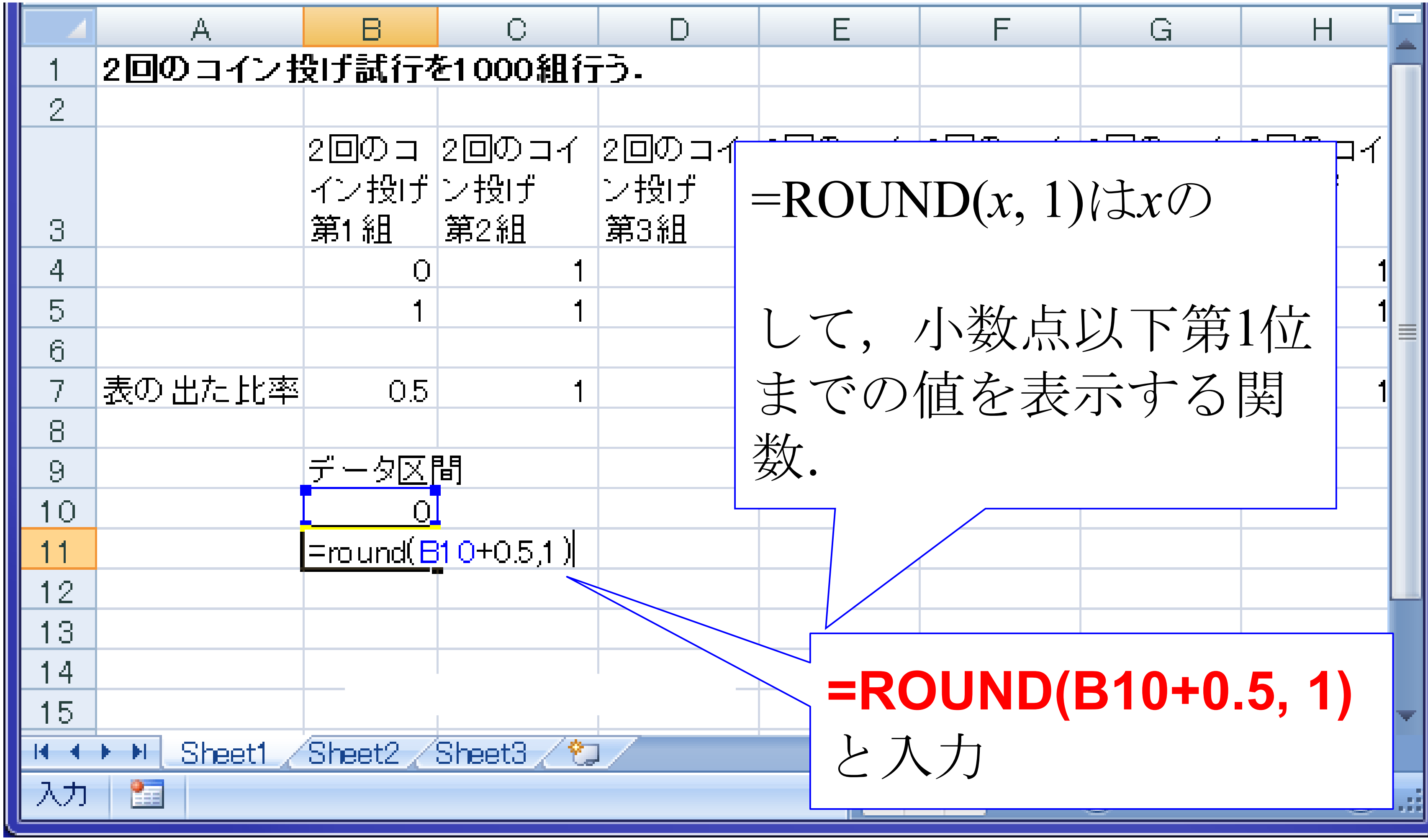


=ROUND(x, 1)はxの

して，小数点以下第1位までの値を表示する関数.

と入力

中心極限定理を体験しよう

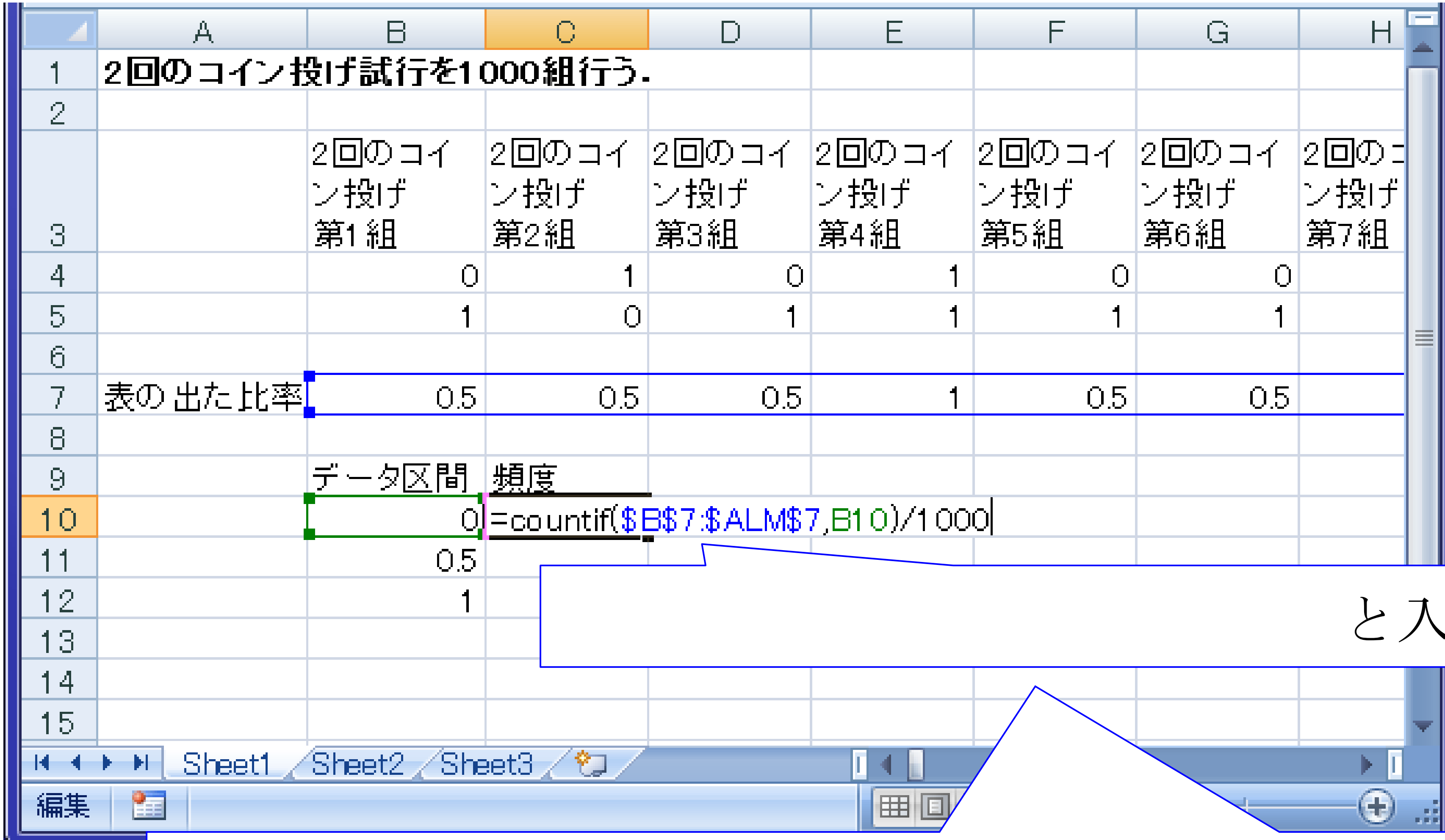


=ROUND(x, 1)はxの

して，小数点以下第1位までの値を表示する関数.

=ROUND(B10+0.5, 1)
と入力

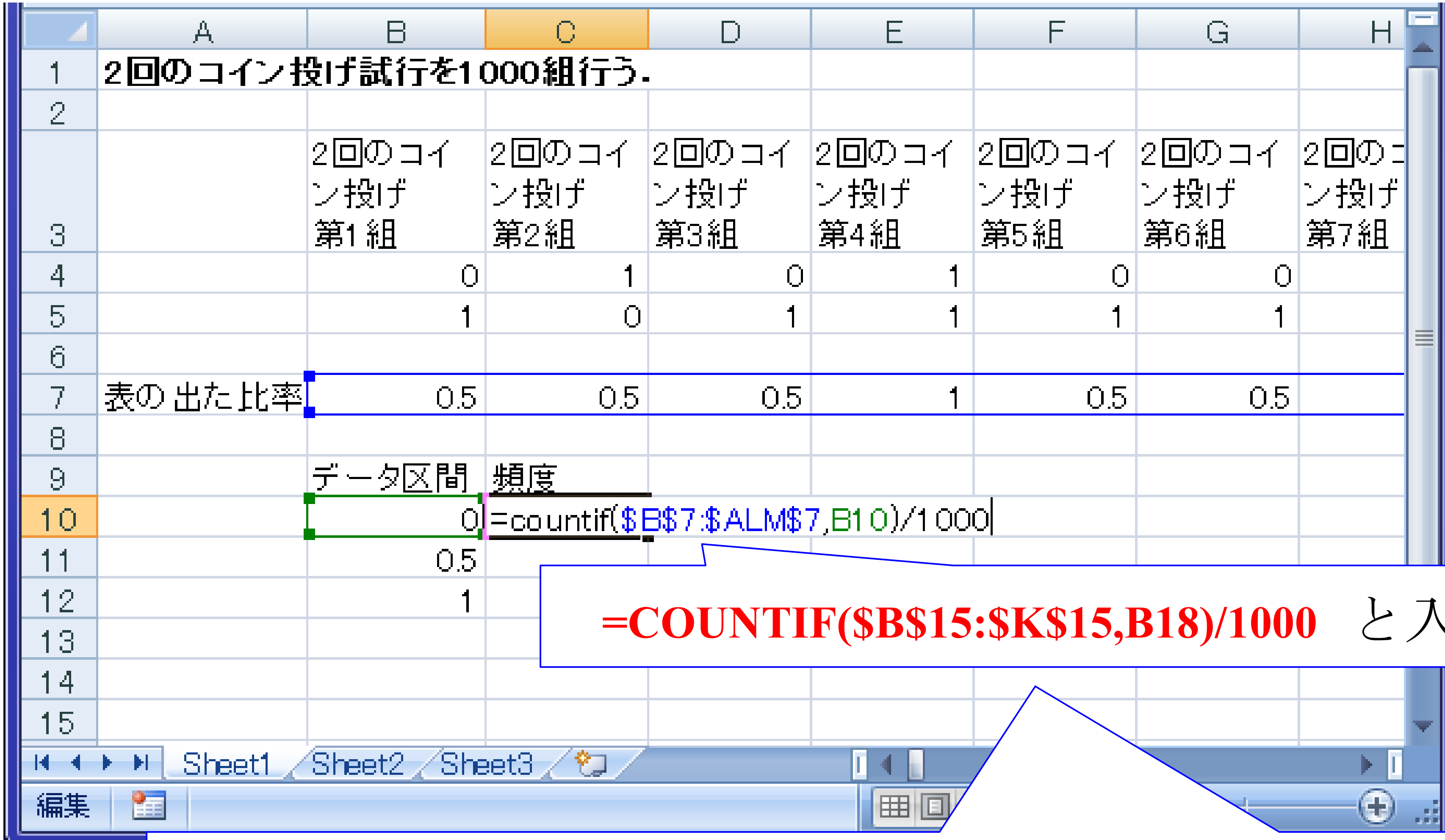
中心極限定理を体験しよう



と入力

1000試行の中で表の出た

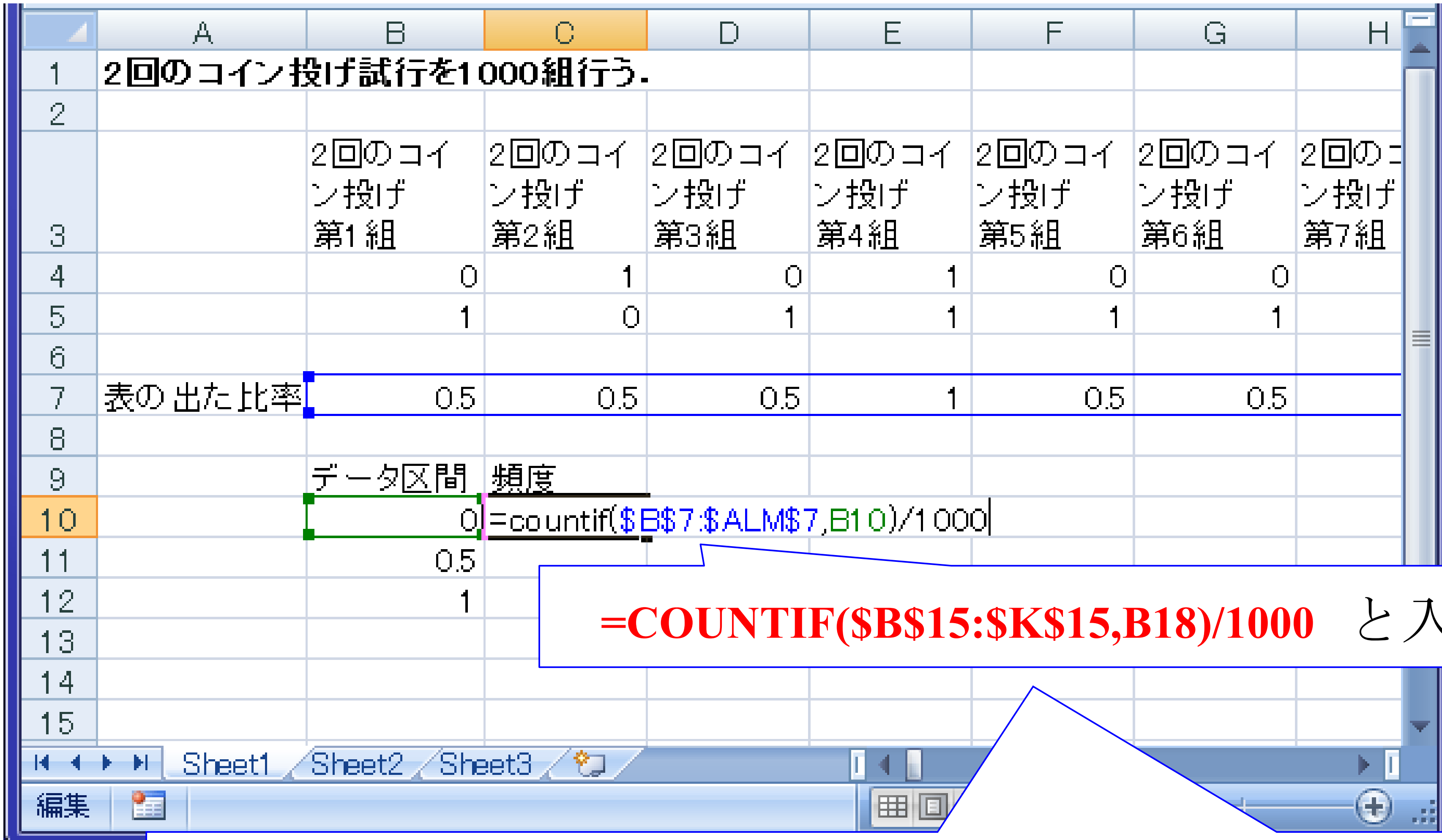
中心極限定理を体験しよう



=COUNTIF($B$15:$K$15,B18)/1000 と入力

1000試行の中で表の出た

中心極限定理を体験しよう

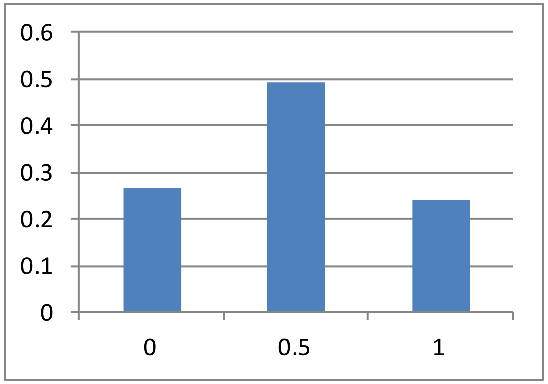


=COUNTIF($B$15:$K$15,B18)/1000　と入力

1000試行の中で表の出た比率が0であった試行回数を求め，さらに1000で割って，その比率を求めている．

中心極限定理を体験しよう

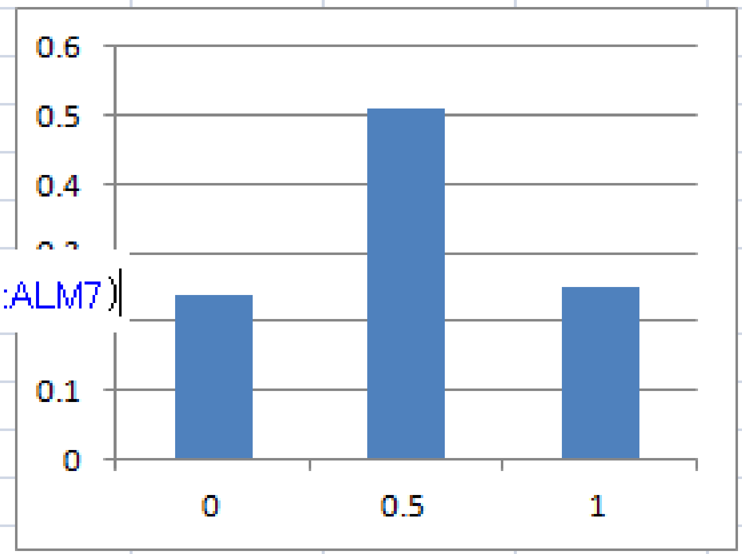
中心極限定理を体験しよう



この分布は，はたして正規分布なのか？

中心極限定理を体験しよう

	A	B	C	D	E	F	G	H
1	2回のコイン投げ試行を1000組行う.							
2								
3		2回のコイン投げ第1組	2回のコイン投げ第2組	2回のコイン投げ第3組	2回のコイン投げ第4組	2回のコイン投げ第5組	2回のコイン投げ第6組	2回のコイン投げ第7組
4		0	0	1	0	1	1	
5		0	1	1	1	1	0	
6								
7	表の出た比率	0	0.5	1	0.5	1	0.5	
8								
9		データ区間	頻度					
10		0	0.239					
11		0.5	0.511					
12		1	0.25					
13								
14		表の出た比率の平均値	=AVERAGE(B7:ALM7)					
15								
16								
17								
18								
19								

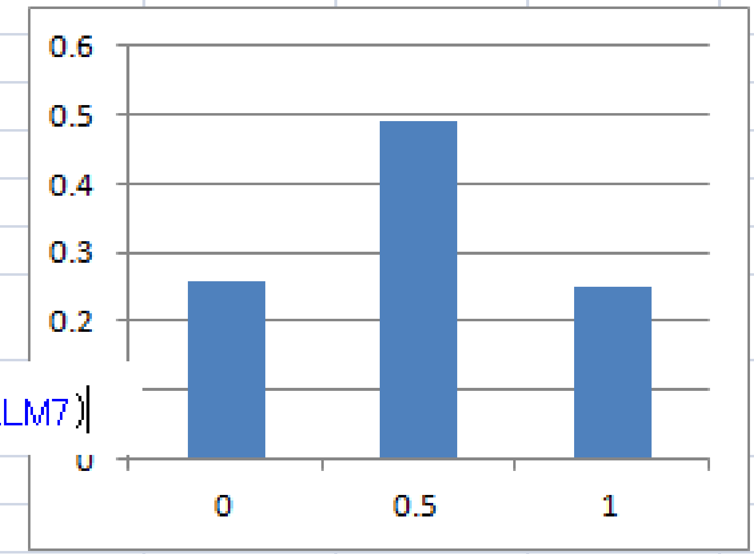


Sheet1 Sheet2 Sheet3

編集 100%

中心極限定理を体験しよう

	A	B	C	D	E	F	G	H
1	2回のコイン投げ試行を1000組行う.							
2								
3		2回のコイン投げ第1組	2回のコイン投げ第2組	2回のコイン投げ第3組	2回のコイン投げ第4組	2回のコイン投げ第5組	2回のコイン投げ第6組	2回のコイン投げ第7組
4		1	0	0	0	1	1	
5		1	1	1	0	1	0	
6								
7	表の出た比率	1	0.5	0.5	0	1	0.5	
8								
9		データ区間	頻度					
10		0	0.259					
11		0.5	0.49					
12		1	0.251					
13								
14		表の出た比率の平均値	0.496					
15		標準偏差推定量	=STDEV(B7:ALM7)					
16								
17								
18								

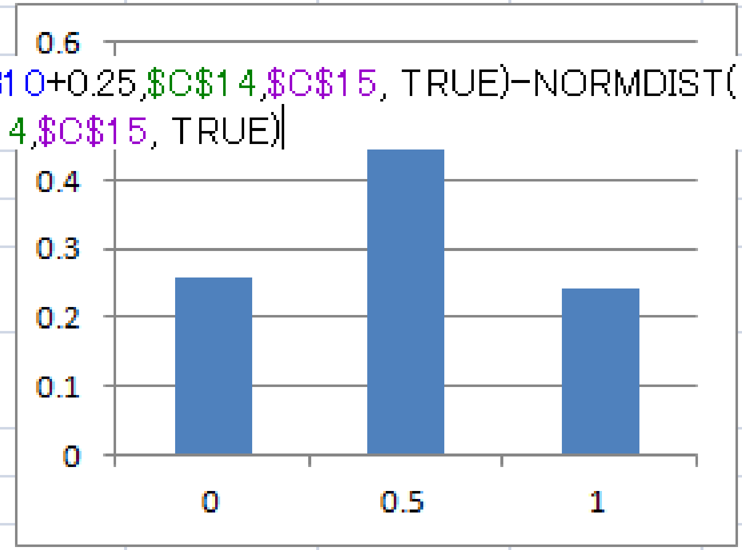


Sheet1 Sheet2 Sheet3

編集 100%

中心極限定理を体験しよう

	A	B	C	D	E	F	G	H	I
1	2回のコイン投げ試行を1000組行う.								
2									
3		2回のコイン投げ第1組	2回のコイン投げ第2組	2回のコイン投げ第3組	2回のコイン投げ第4組	2回のコイン投げ第5組	2回のコイン投げ第6組	2回のコイン投げ第7組	2回のン投け第8組
4		0	0	1	0	0	0	0	
5		0	1	0	0	0	1	1	
6									
7	表の出た比率	0	0.5	0.5	0	0	0.5	0.5	
8									
9		データ区間	頻度	正規分布					
10		0	0.26	=NORMDIST(B10+0.25,$C$14,$C$15, TRUE)-NORMDIST(B10-0.25,$C$14,$C$15, TRUE)					
11		0.5	0.498						
12		1	0.242						
13									
14		表の出た比率の平均値	0.491						
15		標準偏差推定量	0.354323						
16									
17									
18									

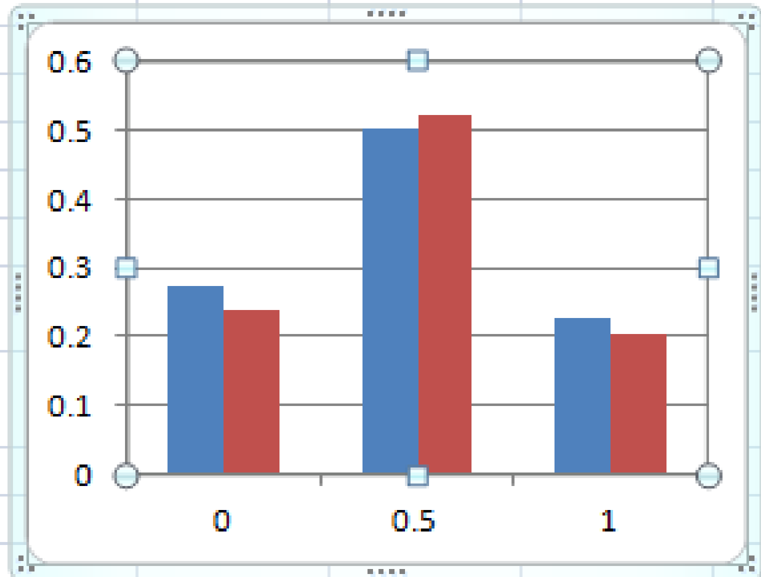


Sheet1 Sheet2 Sheet3

編集 100%

中心極限定理を体験しよう

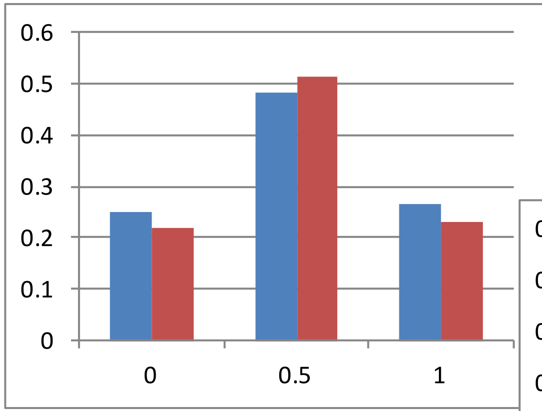
	A	B	C	D	E	F	G	H	I
1	**2回のコイン投げ試行を1000組行う.**								
2									
3		2回のコイン投げ第1組	2回のコイン投げ第2組	2回のコイン投げ第3組	2回のコイン投げ第4組	2回のコイン投げ第5組	2回のコイン投げ第6組	2回のコイン投げ第7組	2回のコイン投げ第8組
4		0	0	0	1	1	0	1	
5		1	1	0	0	1	1	1	
6									
7	表の出た比率	0.5	0.5	0	0.5	1	0.5	1	
8									
9		データ区間	頻度	正規分布					
10		0	0.272	0.24041					
11		0.5	0.502	0.52116					
12		1	0.226	0.205073					
13									
14		表の出た比率の平均値	0.477						
15		標準偏差推定量	0.352271						
16									
17									
18									

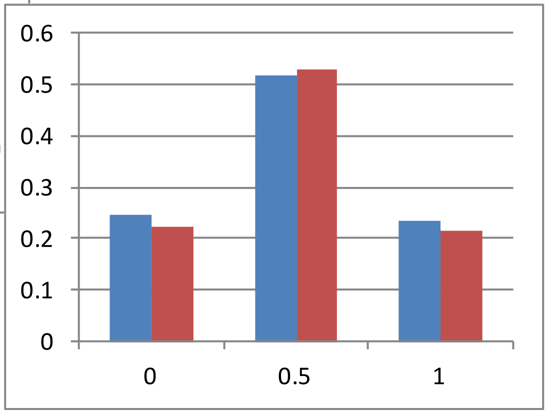


Sheet1 Sheet2 Sheet3

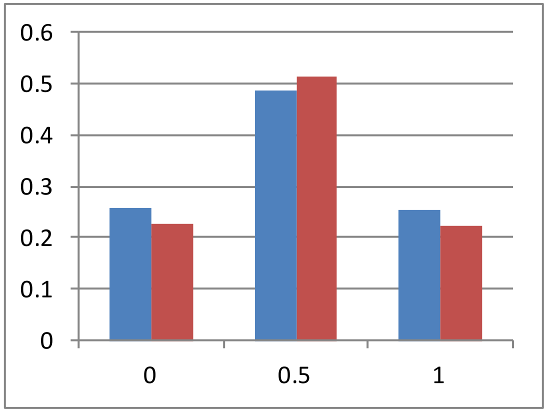
コマンド 平均: 0.322124496 データの個数: 3 合計: 0.966373488 100%

中心極限定理を体験しよう





何度シミュレーションを繰り返しても，分布の差の傾向は変わらない．

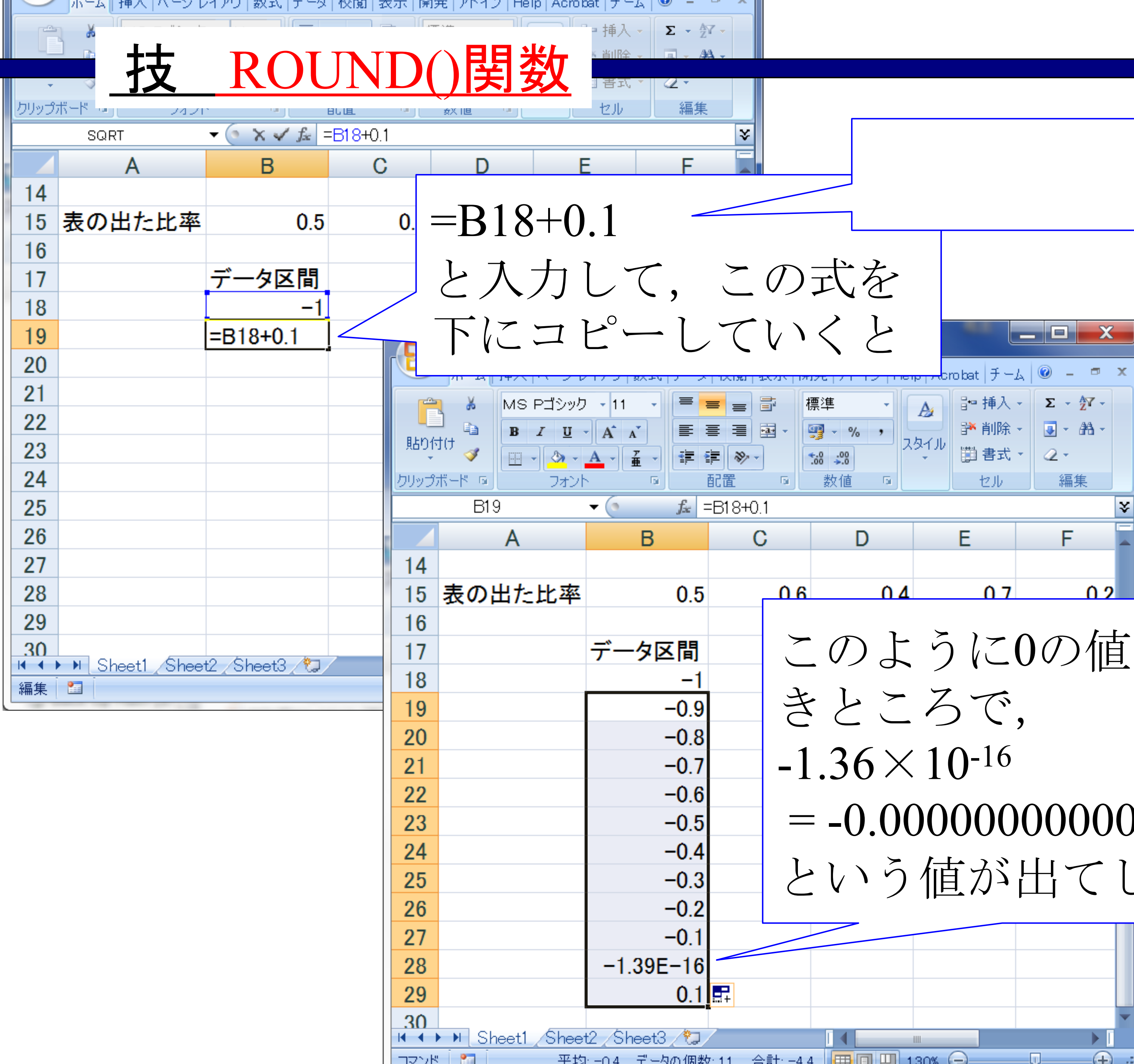


中心極限定理は，もっと極端な分布で試すと，標本の大きさnによる大きな違いを通して体験することができます．
http://www.cmplx.cse.nagoya-u.ac.jp/~furuhashi/education/Statistics_Multivariate/index.html
の第2章に指数分布の例を紹介してあります．Excelファイルも無料でダウンロードできます．よかったら試してみて下さい．

小テスト4.1

10回コインを投げたときに表の出る比率$q/10$を求める．これを1組とする．1000組の結果について表の出る比率の分布を作成し，正規分布と比較せよ．

技　ROUND()関数



=B18+0.1
と入力して，この式を下にコピーしていくと

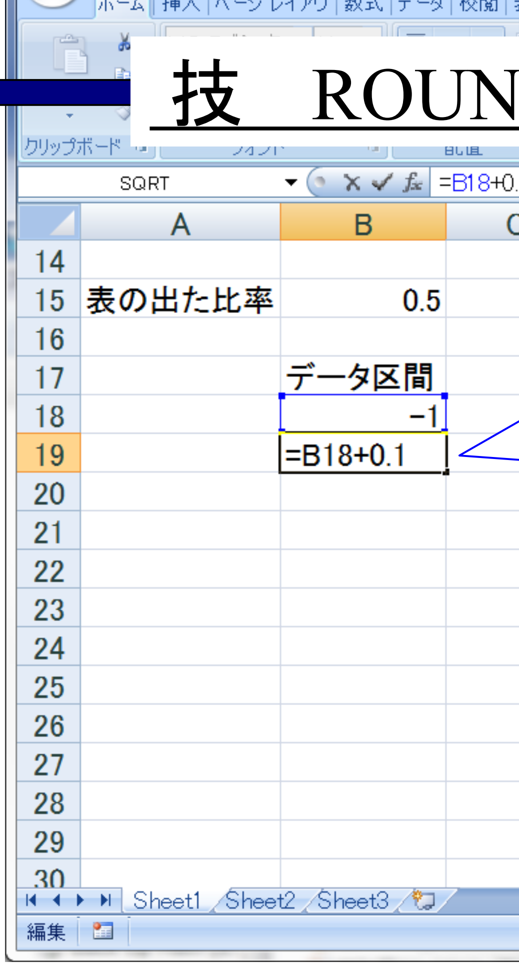
このように0の値が出るべきところで，
-1.36×10^{-16}
＝-0.000000000000000139
という値が出てしまう．

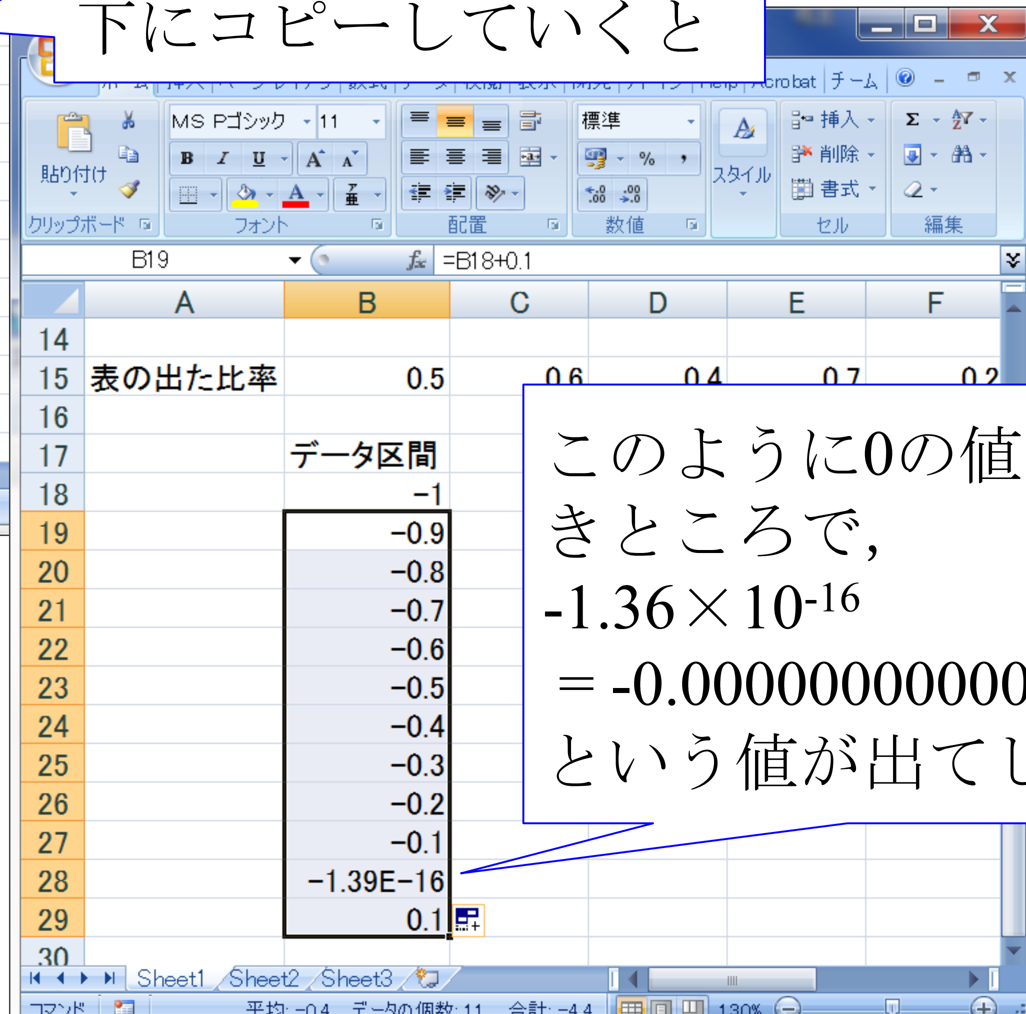
技　ROUND()関数

ROUND() 関数を用いないと

=B18+0.1

と入力して，この式を下にコピーしていくと





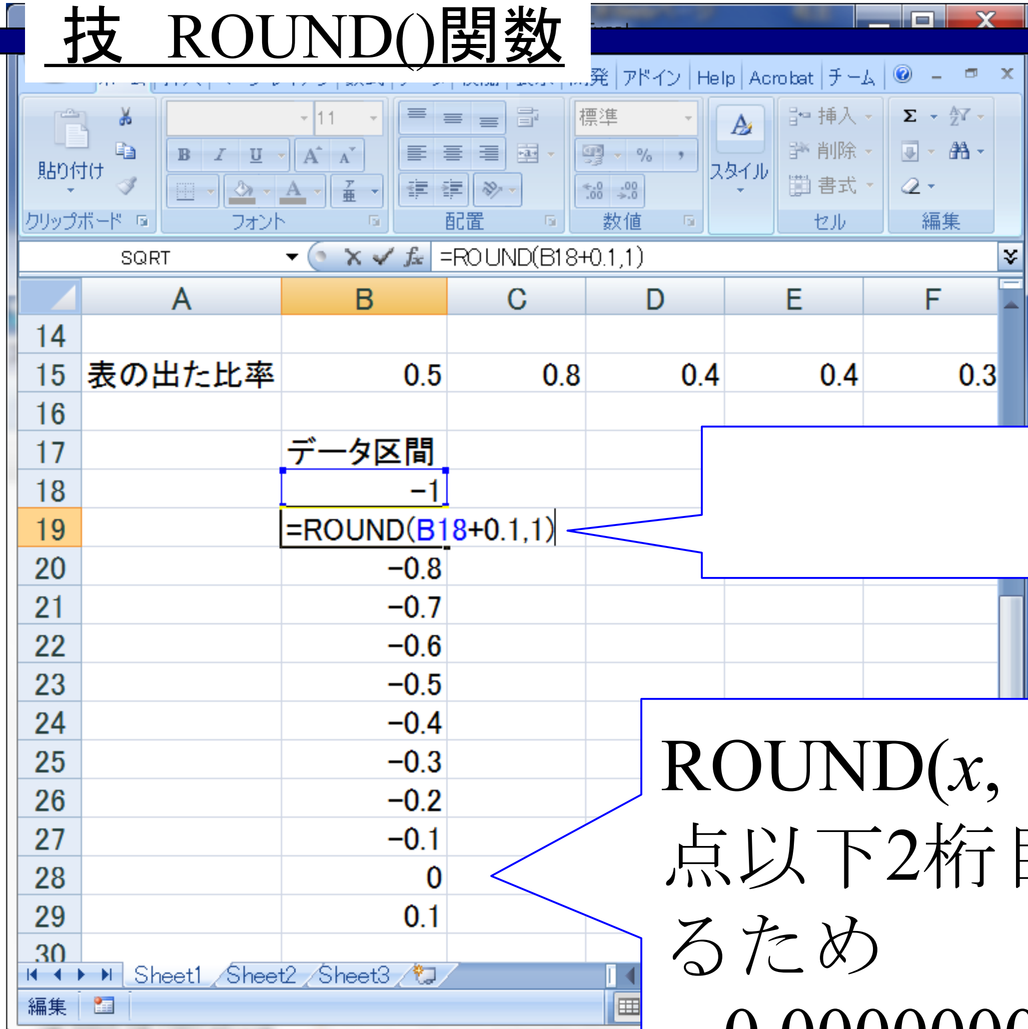
このように0の値が出るべきところで，

-1.36×10^{-16}

$=-0.000000000000000139$

という値が出てしまう．

技 ROUND()関数

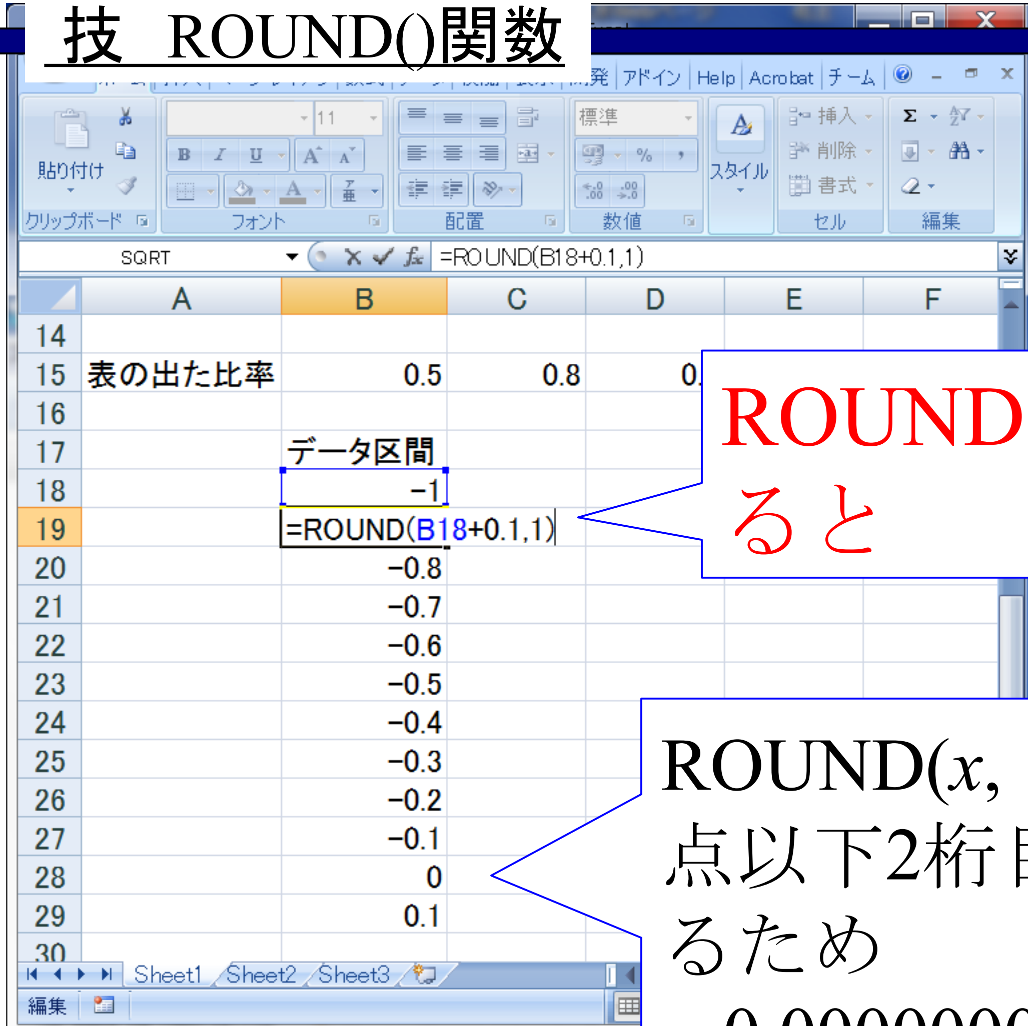


ROUND(x, 1)により，xの小数点以下2桁目が四捨五入されるため

-0.000000000000000139

となる．

技　ROUND()関数



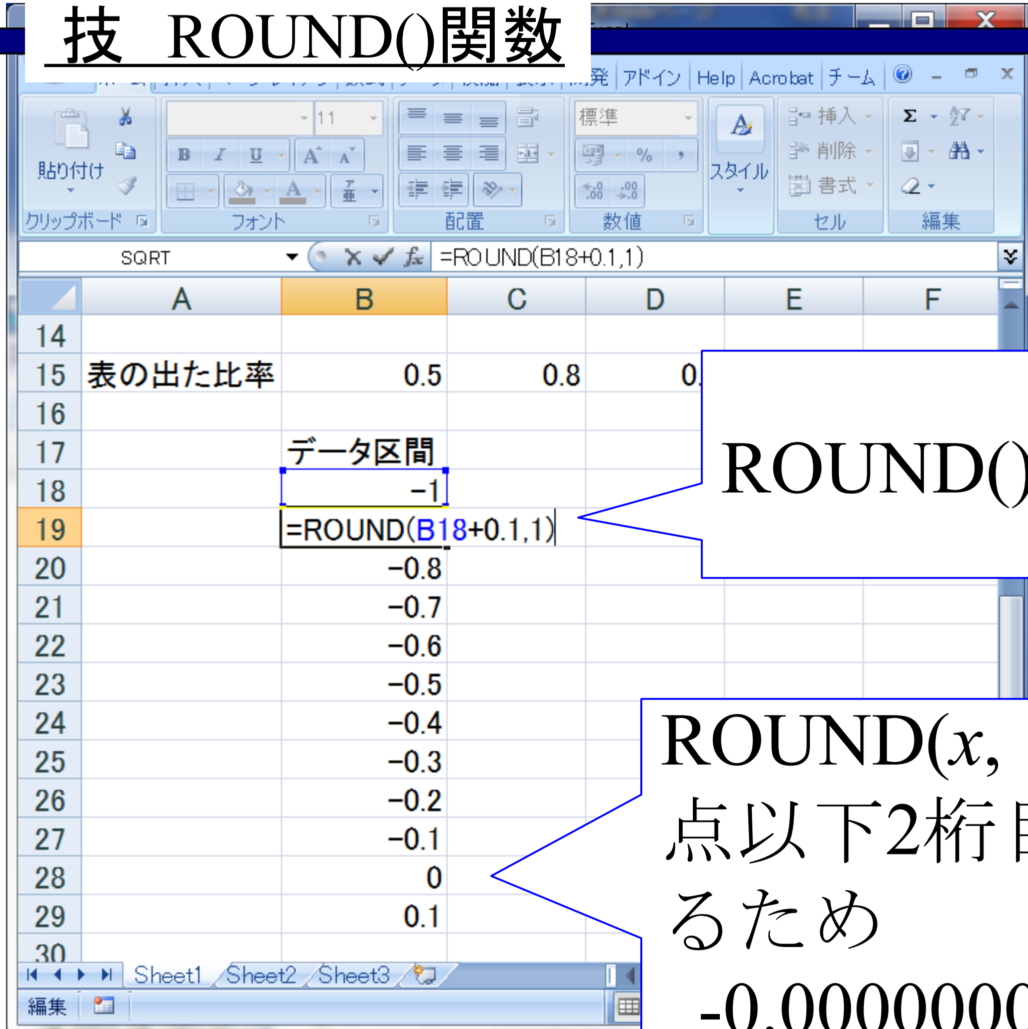
ROUND() 関数を用いると

ROUND(x, 1)により，xの小数点以下2桁目が四捨五入されるため

-0.0000000000000000139

となる．

技　ROUND()関数



ROUND()関数を用いると

ROUND(x, 1)により，xの小数点以下2桁目が四捨五入されるため

-0.000000000000000139

→ 0

となる．

2013年3月

著者：　古橋武
名古屋大学工学研究科計算理工学専攻
furuhashi@cse.nagoya-u.ac.jp