

EXCEL を用いたヒストグラムの作成

「データの分析(統計学1)」 配布資料

経済学部 講師 岩村 英之

1. 準備 EXCEL による計算の基本操作

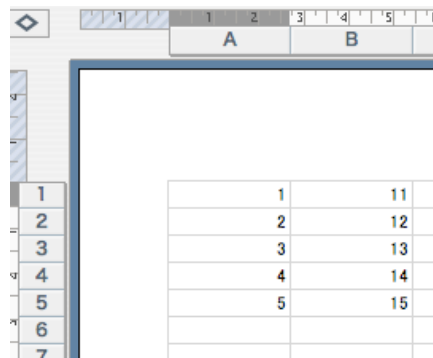
以下では、EXCEL の基本操作として、

- ・式入力による計算
- ・コピー/ペーストを利用した繰り返し計算
- ・「関数」を用いた計算
- ・一部の数値を固定した繰り返し計算

について説明します。実際に EXCEL を起動し、このマニュアル通りに入力し、出力結果を確かめてみてください。

1.1 式入力による計算

まず、A1 セルから B5 セルに下の図のように数値を入力してください。入力する前に、キーボードの「英数」(スペースキーのすぐ左側にある)を押して、半角モードにしてください。EXCEL は、全角で入力された数字を「文字」として扱い、「数字」として扱ってくれません。



1	11
2	12
3	13
4	14
5	15

次に、セル C1 に次のように入力し、enter を押してください。

+A1+B1

すると、「12」という数値がセル C1 に出力されるはずですが、

これは、セル C1 の入力が「セル A1 と B1 を足しなさい」という命令だったため、コンピュータはセル A1 の数値 1 と B1 の数値 11 を足して、12 という数値を出力してきたためです。

命令の最初についている「+」は、「これから入力する内容は‘式’ですよ」というサインです。これをつけなくて、ただ「A1+B1」と入力してしまうと、コンピュータは「文字列」だと判断し、計算せずにそのまま「A1+B1」という文字を出力してきます。

次に、かけ算をやってみましょう。

D1 セルに「+A1*B1」と入力し、enter を押してください。コンピュータはセル A1 とセル B1 のか

け算, すなわち「1×11」を実行し, D1 セルに「11」と出力してくるはずですが. このように, かけ算には「*」(アスタリスク)を用います.

一方, 割り算をさせたいときは「/」(スラッシュ)を使います. E1 セルに「+A1/B1」と入力し, enter を押してください. 「セル A1÷セル B1」すなわち「1÷11」の結果である「0.0909...」が出力されるはずですが.

1.2 コピー/ペーストを利用した繰り返し計算

先ほど, C1 セルに「となりのふたつのセルを合計する」(A1+B1)という命令を入力しました. 同じ計算を C2 から C5 でも実行したい場合, いちいち各セルに「+A2+B2」「+A3+B3」「+A4+B4」...と入力していくのは面倒です.

	A	B	C	D	E
1	1	11	12	11	0.090909091
2	2	12			
3	3	13			
4	4	14			
5	5	15			
6					

そこで, 次のような簡便な方法があります.

まず, セル C1 をコピーしてください(セル C1 にカーソルを移動し, 「編集」—「コピー」).

次に, セル C2 から C5 までを範囲指定し, 「編集」—「ペースト」.

すると, セル C2 から C5 に足し算の結果が出力されるはずですが.

	A	B	C
1	1	11	12
2	2	12	14
3	3	13	16
4	4	14	18
5	5	15	20
6			

これは, 最初の「コピー&ペースト」によって, セル C1 の「右側のふたつのセルを合計しない」という「計算式」が C2 から C5 にコピーされたためです. このように, 計算式が入力されているセルを「コピー&ペースト」すると, 計算結果「12」ではなく, 「右側の2つのセルを合計しろ」という計算式そのものがコピーされます. この性質を利用すれば, 同じ式を繰り返し入力する手間が省けます.

同じように, 先ほどかけ算式を入力した D1 セルをコピーし, D2 から D5 にペーストすれば, かけ算が繰り返されます. 同様に, E1 セルをコピーして E2 から E5 にペーストすれば, 割り算が繰り返されます. これらは各自試してみてください.

1.3 「関数」を用いた計算

セル A6 に「+sum(A1:A5)」と入力し、enter を押してください。

関数 sum(○:△)は、「○から△までの数値を合計せよ」という命令です。従って、ここでは A1 から A5 までの合計である「15」という数値が出力されたはずですが、

次に、セル B6 に「+max(B1:B5)」と入力し、enter を押してください。

また、セル B7 には「+min(B1:B5)」と入力し、enter を押してください。

これらは、「セル B1 から B5 までの中で最大値(max)・最小値(min)を見つけよ」という命令です。したがって、B6 には最大値 15 が、B7 には最小値 11 がそれぞれ出力されるでしょう。

最大値/最小値を見つけるこれらの関数は、データの数が膨大なときには必要不可欠です。

1.4 固定した数値を用いた繰り返し計算

先ほどの操作で、セル A6 には「15」が出力されているはずですが、

ここで、セル F1 に「+A1/A6」と入力し enter を押してください。「1÷15」の答えである「0.0666...」が出力されるはずですが、

次に、セル A2 からセル A5 の数値も、同じようにセル A6 で割りたいとしましょう。このとき、やはり「+A2/A6」「+A3/A6」…「+A5/A6」と次々に入力していくのは面倒です。そこで、たいていの初心者は 1.2 でやったように、セル F1 をコピーして、セル F2 からセル F5 にペーストしてしまいます。しかし、残念ながらセル F2 からセル F5 には下の図のように予想外の結果が出力されてしまうでしょう。

	A	B	C	D	E	F
1	1	6	7	6	0.166666667	0.066666667
2	2	7	9	14	0.285714286	#DIV/0!
3	3	8	11	24	0.375	#DIV/0!
4	4	9	13	36	0.444444444	#DIV/0!
5	5	10	15	50	0.5	#DIV/0!
6	15					

これは、セル F1 をセル F2 にコピーすると、「+A1/A6」が「+A2/A7」になってしまうからです。すなわち、A1 が A2 へとズれるだけでなく、A6 のほうも A7 にズれてしまうのです。同様に、F2 にペーストすると、A1 が A3 へとズれると同時に、A6 のほうも A8 にズれてしまいます。A7 や A8 には何の数値も入力されていないので、EXCEL はこれらを「0」と判断します。したがって、F2 や F3 には「A2÷0」「A3÷0」の結果が出力されてしまうのです。

ペーストしてもセル A6 はズれていかないようにする、つまり A6 を固定したまま繰り返し計算をするためには、最初のセルである F1 に「+A1/\$A\$6」と入力する必要があります。「\$」は、「コピー

—&ペーストしてもセル A6 だけは動かすな」という命令です。

実際に、セル F1 に「+A1/\$A\$6」と入力した後、セル F1 をコピーして F2 から F5 へとペーストしてみてください。確かに、「A2÷A6」、「A3÷A6」、「A4÷A6」、「A5÷A6」の結果が出力されるはずですが、

	A	B	C	D	E	F
1	1	6	7	6	0.166666667	+A1/\$A\$6
2	2	7	9	14	0.285714286	
3	3	8	11	24	0.375	
4	4	9	13	36	0.444444444	
5	5	10	15	50	0.5	
6	15					
7						

	A	B	C	D	E	F
1	1	6	7	6	0.166666667	0.066666667
2	2	7	9	14	0.285714286	0.133333333
3	3	8	11	24	0.375	0.2
4	4	9	13	36	0.444444444	0.266666667
5	5	10	15	50	0.5	0.333333333
6	15					
7						

*この操作は、分散を計算するときに威力を発揮します。

以上で必要最低限の準備は整いましたので、次にヒストグラム作成の手順を説明しましょう。

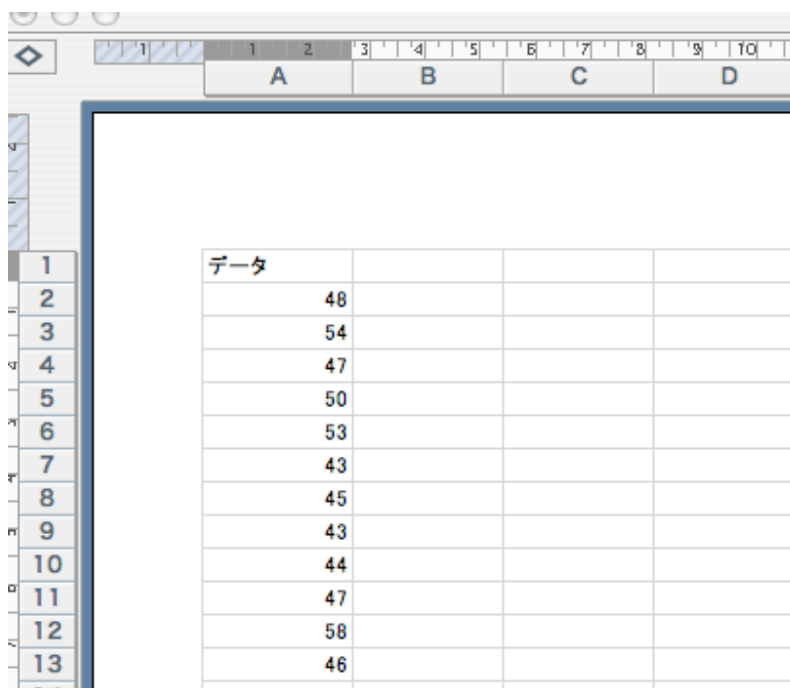
2. ヒストグラムの作成手順

教科書 p.23 の「女子大生の体重」のデータを用いて、EXCEL を利用して度数分布表・ヒストグラムを作成する手順を説明します。

手順1. データを A2 セルから A81 セルまで入力する。

* 私のウェブサイトに入力済みファイル(weight.xls)がアップされていますので、面倒な場合はダウンロードして使ってください。

<http://opinion.nucba.ac.jp/~iwamura/>



	A	B	C	D
1	データ			
2	48			
3	54			
4	47			
5	50			
6	53			
7	43			
8	45			
9	43			
10	44			
11	47			
12	58			
13	46			

手順2. 度数分布表をつくる準備をする。

C1 に「階級下限」、D1 に「階級上限」、E1 に「階級値」、F1 に「度数」、G1 に「相対度数」とそれぞれ入力する。

C	D	E	F	G
階級下限	階級上限	階級値	度数	相対度数

手順3. 各階級の範囲を決め、その下限と上限を C 列と D 列に入力していきます。

教科書によれば、最初の階級は「36kg から 40kg」です。したがって、C2 セルに 36 を、D2 セルに 40 をそれぞれ入力します。以下、2 番目以降の階級についても同様に、その上限と下限を入力していきます。

C	D	
階級下限	階級上限	階級
36	40	
41	45	
46	50	
51	55	
56	60	
61	65	
66	70	

手順4. E 列に階級値を入力します.

階級値は階級の中間の値に設定するのが一般的なので, 各階級の階級値は次の式で計算します.

E2 セルには最初の階級の階級値が入りますので, E2 セルに「 $+(C2+D2)/2$ 」と入力し, enter を押します. これは, 「C2 セルと D2 セルの数値を足して2で割った数値を求めよ」という命令です.

階級下限	階級上限	階級値	度数
36	40	$+(C2+D2)/2$	
41	45		

最初の「+」を忘れないよう気をつけてください.

以下の階級についても同じことを繰り返せばよいので, 準備 1.2 で学んだ方法にならって, たったいま計算した E2 セルをコピーして, E3 から E8 にペーストしてください. こうすると, E2 と同じ計算を, セルを一段ずつずらしながら E3 から E8 についても実行してくれます.

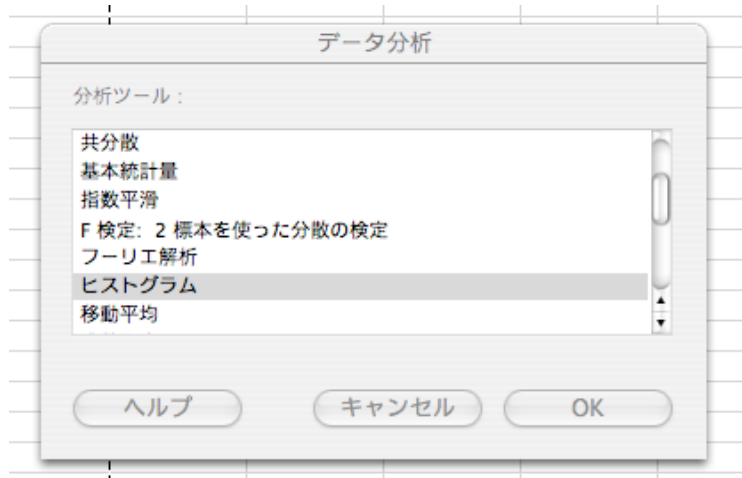
以下のような結果が得られるはずです.

C	D	E	度数
階級下限	階級上限	階級値	
36	40	38	
41	45	43	
46	50	48	
51	55	53	
56	60	58	
61	65	63	
66	70	68	

これで度数分布表を作る用意ができましたので, 実際に度数をカウントしていきましょう.

手順5. 各階級の度数のカウントには、「分析ツール」を用います。

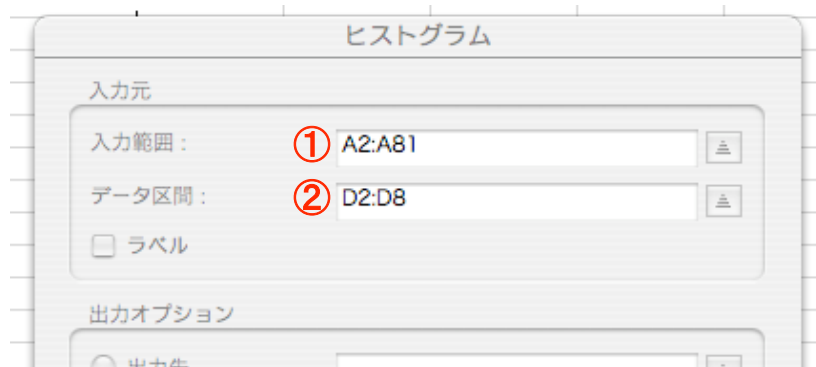
EXCEL の「ツール」メニューから「分析ツール」を選択してください。次のようなボックスが現れますので、「ヒストグラム」を選択し、「OK」をクリックします。



手順6. ヒストグラム画面が現れますので、「入力範囲」と「データ区間」を指定していきま

す。
「入力範囲」とはデータの入力されている範囲のことです。ここでは、A2 から A81 に体重データが入力されていますので、①に「A2:A81」と入力します。

「データ区間」とは、各階級の上限のことです。ここでは D2 から D8 に入力されていますので、②に「D2:D8」と入力します。



「OK」をクリックすれば、EXCEL が度数をカウントし、その結果を新しいワークシートに表示します。結果は次のように表示されます。

	A	B
1	データ区間	頻度
2	40	3
3	45	11
4	50	33
5	55	19
6	60	7
7	65	5
8	70	2
9	次の級	0
10		

手順7. B2 から B8 までをコピーし(B9 まで含めないように注意すること), 最初に作成しておいた表の「度数」の列に貼付けます.

D	E	F	
階級上限	階級値	度数	階級
40	38	3	
45	43	11	
50	48	33	
55	53	19	
60	58	7	
65	63	5	
70	68	2	

手順8. 相対度数を計算します.

最初に F9 セルに「+sum(F2:F8)」と入力し, enter を押してください.

データの合計である 80 が出力されるはずです.

次に, 相対度数を計算するために, F2 から F8 までをデータ数 80(セル F9)で割る必要があります. そこで, はじめに G2 セルに「+F2/\$F\$9」と入力し, enter を押してください.

すると, 1 番目の階級の相対度数が出力されます.

このとき, 「\$F\$9」という具合に「\$」をつけることに気をつけてください. 「\$」が必要な理由は準備 1.4 で説明した通りです.

	F	G
	度数	相対度数
18	3	+F2/\$F\$9
13	11	

F2 セルをコピーし, F3 から F8 までにペーストすると, 2 番目から 7 番目までの階級の相対度数が出力されます.

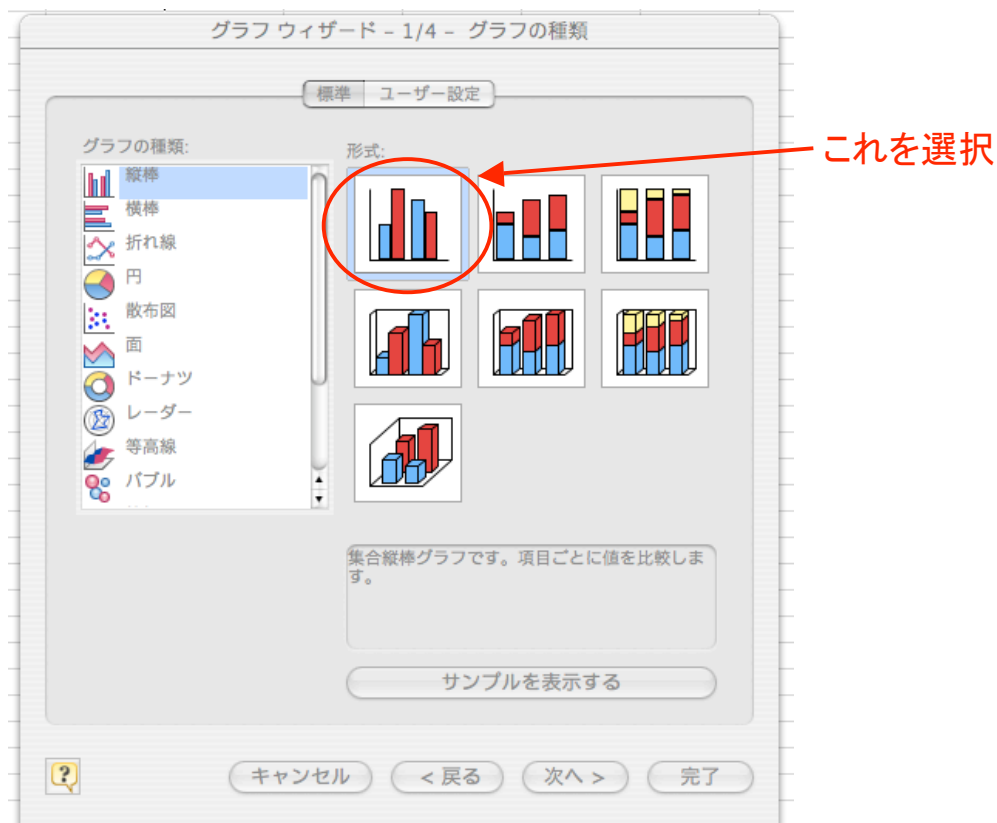
G
相対度数
0.0375
0.1375
0.4125
0.2375
0.0875
0.0625
0.025

以上で度数分布表が完成しました。

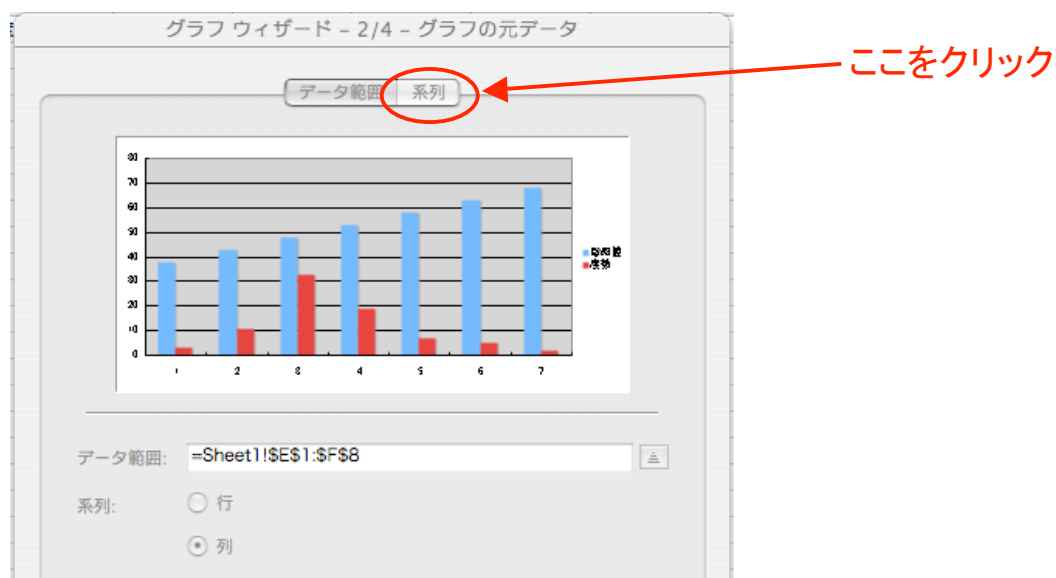
次に、表をグラフ化してヒストグラムを作成する手順に入ります。

手順9. ヒストグラムに必要なデータは、階級値と度数のみです。

従って、階級値と度数が入力されているE1 からF8までを範囲指定し、「挿入」メニューから「グラフ」を選択してください。下のようなグラフ作成画面が現れますので、「形式」のところから左上のグラフを選択し、「次へ」をクリックしてください。



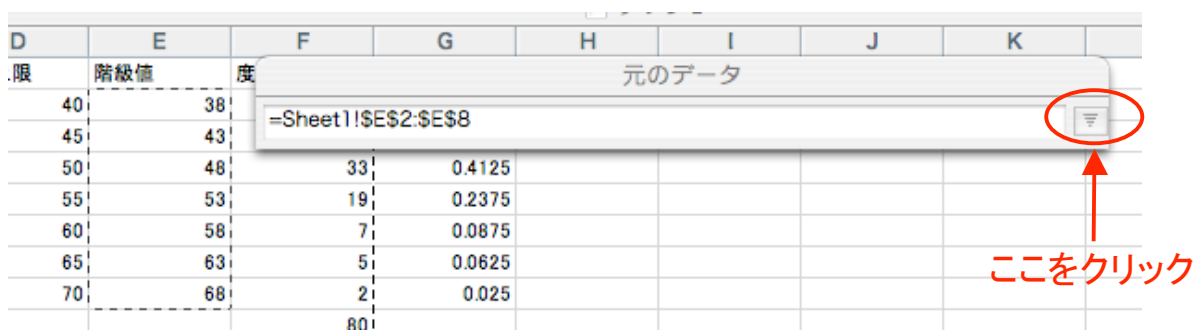
次のようなグラフが表示されますが、これは私たちの求めるものではありません。
調整するために「系列」をクリックしてください。



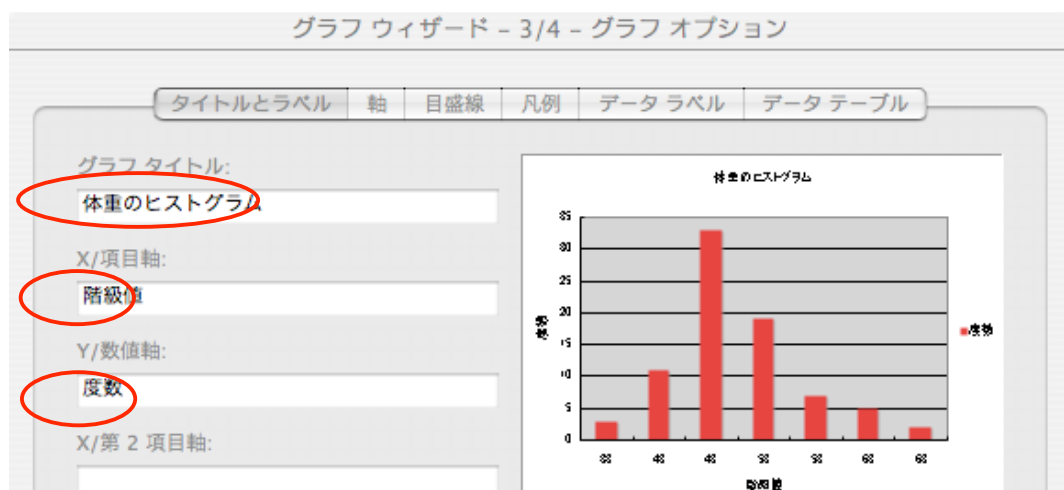
「階級値」(下図①)を選択し、その下の「削除」(下図②)ボタンをクリックしてください。
また、「項目軸ラベルに使用」のところの右端のボタン(下図③)をクリックしてください。



次のようなウィンドウが現れますので、E2 から E8 までを範囲指定し、右端のボタンをクリックして元の画面に戻ってください。戻ったら「次へ」をクリック。

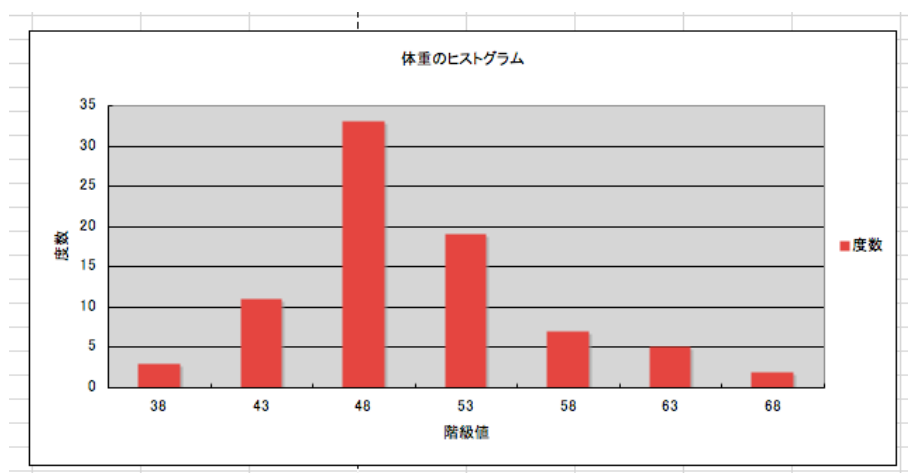


手順 10. 「タイトルとラベル」をクリックし、適当なグラフタイトル、X 軸の名前、Y 軸の名前を入力してください。この例では、グラフタイトルを「体重のヒストグラム」、X 軸・Y 軸の名前をそれぞれ「階級値」・「度数」としています。



「次へ」をクリックすると、ワークシートにグラフが出力されます。

これでヒストグラムが出来上がりました。非常に「きれいな」ヒストグラムになっていることを確認してください。



(注)

名古屋商科大学では Mac ユーザがほとんどだと思いますので、以上の説明には Mac 版 EXCEL の画面を用いています。とは言え、この程度の作業あれば、Windows ユーザの私もほとんど戸惑うことはありませんでした。したがって、Windows 版 EXCEL の画面も大差はなく、この資料の内容がほぼそのまま通用します。レポートは Windows で作成しても構いません。

参考文献

縄田和満(2000), 『EXCEL による統計入門』(第2版), 朝倉書店.

特に第6章を参考にしました.