

EXCEL を用いたヒストグラムの作成

「2009 年度 基礎文献購読」配布資料
明治学院大学 国際学部 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。



	A	B	C	D	E	F
1	1	11				
2	2	12				
3	3	13				
4	4	14				
5	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。

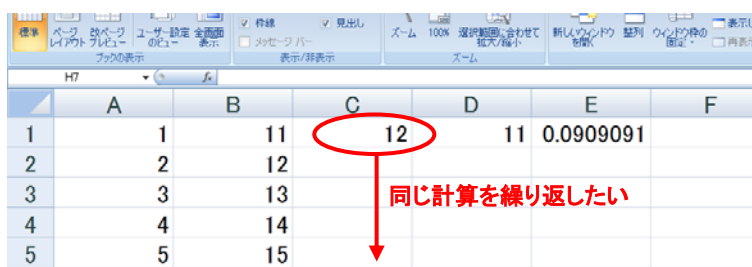
次に、かけ算をやってみましょう。

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。



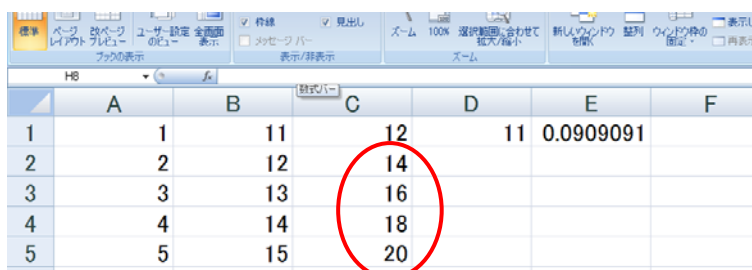
	A	B	C	D	E	F
1	1	11	12	11	0.0909091	
2	2	12				
3	3	13				
4	4	14				
5	5	15				

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	D	E	F
1	1	11	12	11	0.0909091	
2	2	12	14			
3	3	13	16			
4	4	14	18			
5	5	15	20			

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さい」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	G
1	1	11	12	11	0.0909091	0.0666667	
2	2	12	14	24	0.1666667	#DIV/0!	
3	3	13	16	39	0.2307692	#DIV/0!	
4	4	14	18	56	0.2857143	#DIV/0!	
5	5	15	20	75	0.3333333	#DIV/0!	
6	15	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、

The top screenshot shows the Excel formula bar with the formula $=A1/\$A\6 entered in cell G1. The spreadsheet data is as follows:

	A	B	C	D	E	F	G
1	1	11	12	11	0.0909091	0.0666667	$=A1/\$A\6
2	2	12	14	24	0.1666667	#DIV/0!	
3	3	13	16	39	0.2307692	#DIV/0!	
4	4	14	18	56	0.2857143	#DIV/0!	
5	5	15	20	75	0.3333333	#DIV/0!	
6	15	15					

The bottom screenshot shows the calculated results for the same data, with the formula bar set to H6:

	A	B	C	D	E	F	G
1	1	11	12	11	0.0909091	0.0666667	0.0666667
2	2	12	14	24	0.1666667	#DIV/0!	0.1333333
3	3	13	16	39	0.2307692	#DIV/0!	0.2
4	4	14	18	56	0.2857143	#DIV/0!	0.2666667
5	5	15	20	75	0.3333333	#DIV/0!	0.3333333
6	15	15					
7		11					

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. 教科書にある体重データを 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://www1.meijigakuin.ac.jp/~iwamura/>

	A	B	C	D	E	F	G	H	I	J
1	体重(kg)									
2	48									
3	54									
4	47									
5	50									
6	53									
7	43									
8	45									
9	43									
10	44									
11	47									
12	58									
13	46									
14	46									
15	63									
16	49									

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.

B	C	D	E	F	G	H	I	J
	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.

B	C	D	E	F	G	H	I	J	K
	階級下限	階級上限	階級値	度数	相対度数				
	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.

B	C	D	E	F	G	H
	階級下限	階級上限	階級値	度数	相対度数	
	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が.

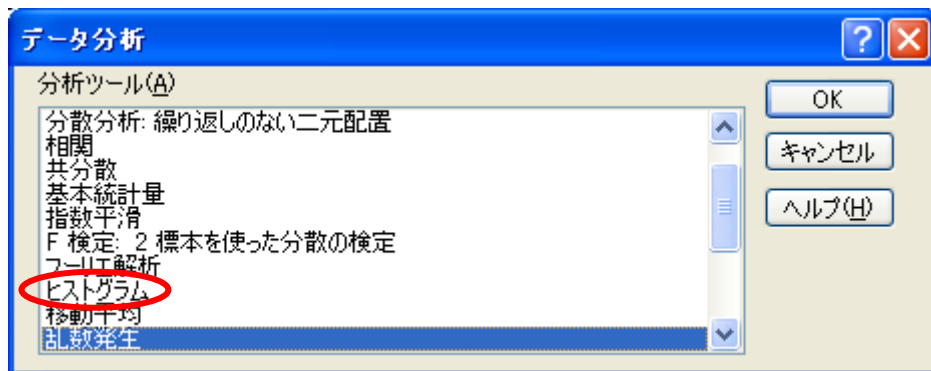
B	C	D	E	F	G	H
	階級下限	階級上限	階級値	度数	相対度数	
	36	40	38			
	41	45	43			
	46	50	48			
	51	55	53			
	56	60	58			
	61	65	63			
	66	70	68			

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「データ」メニューから「データ分析」を選択してください。



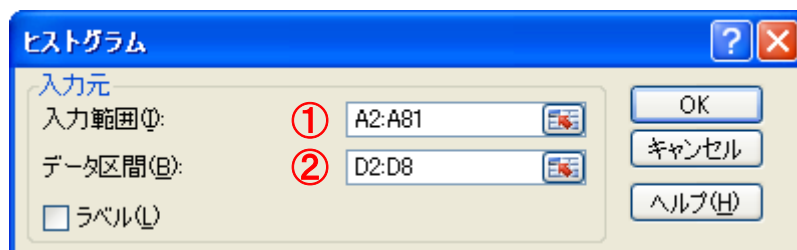
次のようなボックスが現れますので、「ヒストグラム」を選択し、「OK」をクリックします。



- 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。「頻度」と書いてあるところに度数が出力されています。

	A	B	C	D	E
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。

	C	D	E	F	G	H	I
1	階級下限	階級上限	階級値	度数	相対度数		
2	36	40	38	3			
3	41	45	43	11			
4	46	50	48	33			
5	51	55	53	19			
6	56	60	58	7			
7	61	65	63	5			
8	66	70	68	2			
9							
10							

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が。

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。

D	E	F	G	H	I
階級上限	階級値	度数	相対度数		
40	38	3	+F2/\$F\$9		
45	43	11			
50	48	33			
55	53	19			
60	58	7			
--	--	--			

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。

C	D	E	F	G	H	I
階級下限	階級上限	階級値	度数	相対度数		
36	40	38	3	0.0375		
41	45	43	11	0.1375		
46	50	48	33	0.4125		
51	55	53	19	0.2375		
56	60	58	7	0.0875		
61	65	63	5	0.0625		
66	70	68	2	0.025		
			80			

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。

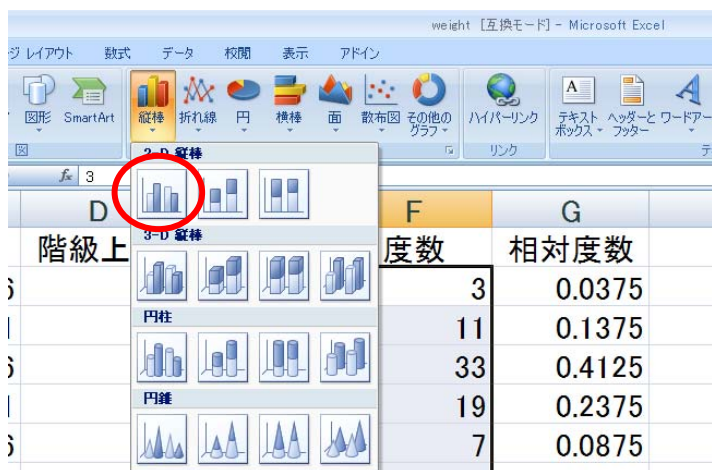
手順9. ヒストグラムに必要なデータは、階級値と度数のみです。

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「縦棒」を選択してください。

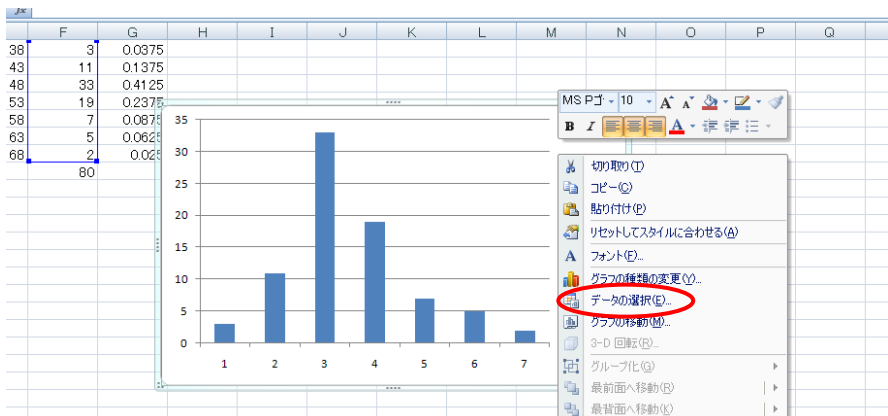
The screenshot shows the Excel ribbon with the '挿入' (Insert) tab selected. The '柱状' (Bar) chart icon is highlighted with a red circle. Below the ribbon, a table of data is shown with columns C through H. A red bracket highlights the range E1:F8, with the text '範囲指定' (Range specified) written in red next to it.

C	D	E	F	G	H
階級下限	階級上限	階級値	度数	相対度数	
36	40	38	3	0.0375	
41	45	43	11	0.1375	
46	50	48	33	0.4125	
51	55	53	19	0.2375	
56	60	58	7	0.0875	
61	65	63	5	0.0625	
66	70	68	2	0.025	
			80		

下のように、縦棒グラフのさらに細かい種類が現れますので、左上のもっともシンプルなものを選んでください。



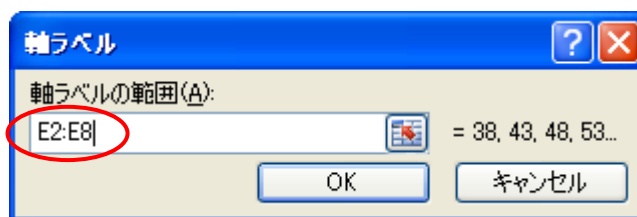
下のようなグラフが現れますが、横軸が階級値になっていません。これを直すため、ポインタをグラフの上にもってきて左クリックし、現れたメニューから「データの選択」を選択してください。



下のようなボックスが現れますので、「横軸ラベル」の「編集」をクリックしてください。

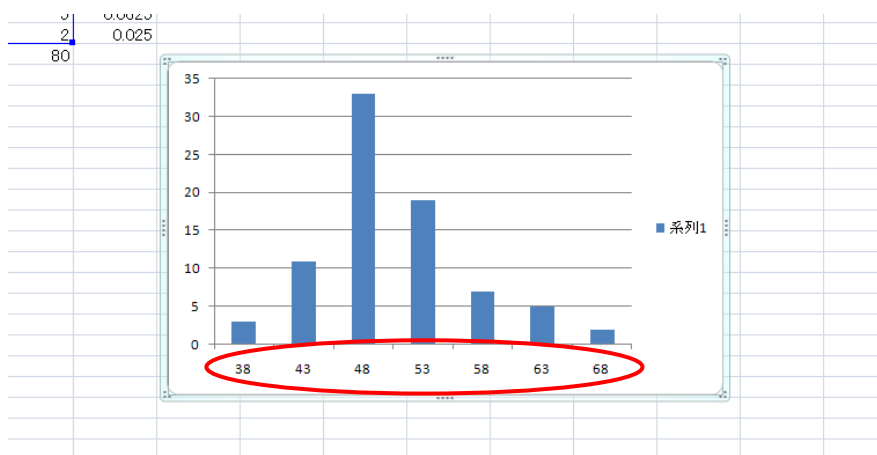


次のようなボックスが現れます。私たちは階級値を横軸に指定したいのですから、階級値の入っているセル、E2 から E8 までを指定して、「OK」をクリックします。



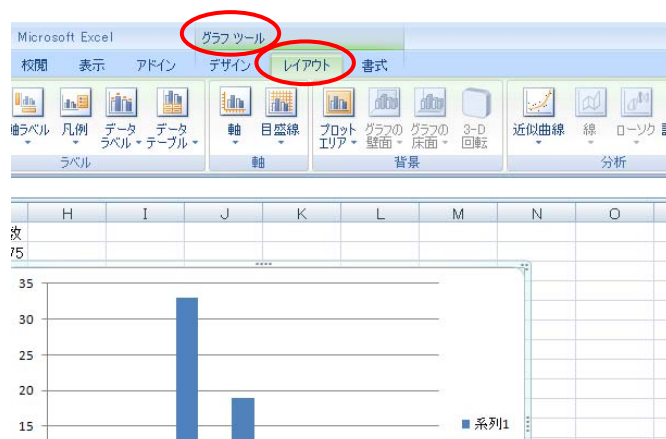
元の画面に戻りますので、「OK」をクリックしてください。

横軸が階級値になっていることを確認してください。



手順 10. グラフにはタイトルがあると親切です。一方、このグラフの場合「凡例」は不要です。次にこれら进行处理しましょう。

グラフをクリックしてアクティブにすると、メニューに「グラフツール」というものが現れます。その中の「レイアウト」を選択してください。



グラフのレイアウトに関する様々なメニューが出てきますので、タイトルや凡例など見やすくなるよう工夫してみてください。

下は、タイトルを加えて凡例を取り除いた例です。



(注)

大学の PC にインストールされている OFFICE のバージョンが 2007 であるため、以上の説明には EXCEL2007 の画面を用いています。ただ、皆さんの中には EXCEL2003 を使っている人も多いと思います。実は、私も 2003 ユーザでしたが、このマニュアルを作成するために 2007 を使い始めました。ですが、2003 を用いて同様の作業を行う場合も基本的な発想は同じなので、さほど苦労しませんでした。とは言え、2003 と 2007 とでは画面がかなり異なってくるため戸惑うこともあるでしょう。2003 ユーザの方は、かつて私が作成したマニュアル(画面は mac ですが windows もほとんど同じです)のほうも参考にしてください。ファイルは私のウェブサイトにあります。

<http://www1.meijigakuin.ac.jp/~iwamura/>

むろんレポートは 2007, 2003 どちら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.

*現在は第 3 版が出ています. 第 3 版は 2007 に対応しています.