

セルの参照・乱数と平均値、サイコロの目

2009/6/8 長島雅裕

1. 少し高度なセルの参照の仕方

掛け算の九九の表を、表計算によって作成してみよう。九九とは以下のようなものである。

	A	B	C	D	E	F	G	H	I	J
1		1	2	3	4	5	6	7	8	9
2	1	1x1	2x1	3x1	4x1	5x1	6x1	7x1	8x1	9x1
3	2	1x2	2x2	3x2	4x2	5x2	6x2	7x2	8x2	9x2
4	3	1x3	2x3	3x3	4x3	5x3	6x3	7x3	8x3	9x3
5	4	1x4	2x4	3x4	4x4	5x4	6x4	7x4	8x4	9x4
6	5	1x5	2x5	3x5	4x5	5x5	6x5	7x5	8x5	9x5
7	6	1x6	2x6	3x6	4x6	5x6	6x6	7x6	8x6	9x6
8	7	1x7	2x7	3x7	4x7	5x7	6x7	7x7	8x7	9x7
9	8	1x8	2x8	3x8	4x8	5x8	6x8	7x8	8x8	9x8
10	9	1x9	2x9	3x9	4x9	5x9	6x9	7x9	8x9	9x9

まず、A列と1行目に、掛け算に必要な1～9までの数字を書いておこう。これも最初に1だけ入力し、右下の黒四角をドラッグして「連続データ」にすれば、すぐに入力できる。

次に、B2セルで計算をしよう。このセルでは、B1セルの値とA2セルの値を書ければよいので、 $=B1*A2$ とすれば良さそうである。試しに、これでB2セルの右下の黒四角をドラッグし、B10セルまで持っていくことで、「1の段」を計算してみよう。

すると、意図した通りにはなっていないことに気付くだろう。試しにB10セルの中身を見てみると、 $=B9*A10$ となっているはずである。これは、最初のB2セルでの計算が、「一つ上のセル*一つ左のセル」と解釈されてしまったためだ。九九にするためには、「一つ上のセル」ではなくて、「常に1列目のセル」を参照してもらわなければならない。

このための方法がある。「\$」をつけるのである。B2セルの内容を、 $=B$1*A2$ としてみよう。このセルの値は変わらないはず。これを下までドラッグすると、例えばB10セルの中身は $=B$1*A10$ になり、九九の1の段になっていることがわかる。B1の”1”の部分、強制的に固定したのである。

さて、1の段をまとめて選択し、右のほうへドラッグすれば、ただちに九九が完成しそうであるがそうはいかない。今度は、2の段、3の段、…ではA列の値を参照したいのに、ここでも「一つ左のセル」の値が参照されているからである。これを避けるためには、最初にB2セルで $=B$1*$A2$ としておけばよい。これで、1行目とA列目というのを固定することができ、簡単に九九の表を完成させることができる。

このテクニックは、今後、少し複雑なことをする際に有用になる。

課題:19x19の掛け算の表を作ってみよう。また列の幅も短くして、見やすくしよう(変えたい列を選び、右クリックで「列の幅」を選ぶ)。最後に印刷して提出すること(番号・名前を忘れずに)。

2. 乱数と平均値

サイコロをふると、正しく作られたサイコロならば、1～6の目が等確率で出てくる(確率はいくらでしょう?)。このように、ランダムに発生する数値(の列;数列)のことを「乱数」と呼ぶ。実際に乱数が必要になる場合も多々あるが、統計学を理解する際にも乱数を理解することは大変重要である。

コンピュータはある一定のアルゴリズムに基づいて動いているため、「本当の」乱数を発生させることは不可能である。しかし、「ほぼ」乱数に見える数列を吐き出すプログラムの研究がすすみ、エクセルには乱数を発生させる関数が組み込まれている。rand関数である。

関数であるのでカッコ()が必要だが、引数はとらない。セル内で、 $=rand()$ とすると、区間[0,1)の1様乱数が発生する(0以上1未満でどれも等確率。1は含まれない)。実際に打ち込んでみよう。1未満の小数が書かれるはずである。またファンクションキーF9を押すと、新たな乱数が振り直される。何度か押して、[0,1)の乱数になっていることを体感してください。

次に、平均値を求めてみよう。乱数列[0,1)の平均値は、ちょうど中間の0.5である。実際にそうになっているだろう

うか？

まず、10個の乱数列を発生させよう。例えばA1に=rand()と書き、それをA10までコピーすると、10個の乱数が発生する(F9を押して、色々変えてみてください)。この10個の数の平均値を求めてみる。

セルA11に平均値を出そう。やり方はわかりますね？オートSUMボタンの右側の三角をクリックし、平均値を選ぶ。平均を取る範囲がA1～A10になっていることを確認し、決定する。値は幾らになっただろうか？

0.5付近の人もいれば、0.4以下だったり0.6以上だったりした人もいるだろう。では、さらに何回かF9を押してみよう。0.5のまわりでフラフラとするはず。どれくらいの範囲でフラフラするかを目安を「標準偏差」とか「分散」と言います。これらの定義は次回の講義でします。

さて、無限に乱数を発生させれば、平均値は0.5に近づいていくはずである。それを見てみよう。

左端に1列挿入し、A1～A10まで順に1～10の数字を入れていく。これは何回乱数を発生させたかの数になっている。次に平均値をC列に書いていく。C1セルはB1:B1の平均、C2セルはB1:B2セルの平均、CnセルはB1:Bnセルの平均、とすると、乱数をふるごとに平均値がどう変化するかがわかるはずである。ここで、C1セルに=average(B\$1:B1)と”\$”をつけておくと、C1セルを下にコピーしていても、自動的にB1セルからの平均値を取ってくれる。試しにC10セルの中身を見ると、=average(B\$1:B10)となっているはずである。

ここで、A列を横軸に、C列を縦軸にしたグラフを描いてみよう。乱数を発生し続けると、平均値がどう変化するかが一目でわかるようにするためである。折れ線グラフにして、グラフが煩雑になるのを避けるため、点のマーカーはないものを選んでおこう(詳細は前回のレジュメを参考のこと)。描いたら、何度かF9を押して、グラフの変化を味わってみよう。なんとなく、サイコロをふる回数が増えるほど、0.5に近づくように見えるような気がしないだろうか？

もっとはっきり見るために、100回乱数を発生させてみよう。まず、A1～C10セルまで選択しておいて、右下をドラッグして100行目まで持っていく。これで、B100セルの値は100回目の乱数、C100セルの値は100回目までの平均値、ということになる。次にグラフを変更する。グラフの中を右クリックし、「元のデータ」を選ぼう。今までと同様にマウスで範囲を選択してもいいのだが、ここではもっと簡単に、「値」と「項目軸ラベルに使用」の中の=Sheet1!\$C\$1:\$C\$10の最後の「10」を「100」に変えよう。この意味は、「値」は、「このシートのC1からC100まで」ということだ。両方100に変更すれば、それで100行目までが範囲になる。

図を眺めてみよう。明らかに、回数が増えるほど、平均値は0.5に近づいていっているのがわかるだろう。つまり、理論的に確率分布がわかっても、その確率分布から予想される平均値に近づくには、十分多くのサンプルが必要、ということだ。この意味は統計解析で重要になるので、次回、ゆっくり実習を行う。

課題: 上で描いた100回の乱数についてのグラフを印刷して提出すること。

3. サイコロ

最後にサイコロの目を出してみよう。今まで使ったrand関数は[0,1)区間で一様な乱数を発生させるものであった。これを少し変形すると、サイコロにすることができる。

まず、サイコロの目は、1～6までの整数であることに注意しよう。そこで、方針としては、

- (1)[0,6)区間で一様乱数を発生(0以上6未満)→rand関数
- (2)小数点以下を切り捨て、{0,1,2,3,4,5}をランダムに発生→int関数
- (3)1足して、{1,2,3,4,5,6}にする

これを実現するには、=int(rand()*6)+1とすればよい。rand()*6で、[0,1)区間の乱数が6倍され、[0,6)区間の一様乱数になる。次にint()を作用させることで、小数点以下を切り捨てる(int関数は小数点以下を切り捨て、整数に変換する関数)。最後に+1することで、1～6までの整数をランダムに発生させることができる。

何度かF9を押して、1～6の目が出ることを確かめてみよう。

また2でやったように、沢山のセルでサイコロの目を発生させ、同じくらいの確率で出ているかどうか、見てみよう。

次回はヒストグラムを作り、サイコロの目それぞれが等確率で出ているか視覚的にチェックするところから始めます。

課題(再掲) それぞれ番号・名前を忘れずに

- (1)19x19の掛け算の表を作ってみよう。また列の幅も短くして、見やすくしよう(変えたい列を選び、右クリックで「列の幅」を選ぶ)。最後に印刷して提出すること。
- (2)上で描いた100回の乱数についてのグラフを印刷して提出すること。