

百均の電卓で解ける 統計入門

数学的な厳密さよりも直感的に理解することに
重点を置いています・・・

連載にあたって

これから統計を学ぼうとしている方、忘れてしまったので思い出そうとしている方、あるいは何度もチャレンジしてはみたが、なかなか理解できなかつた方を対象として連載します。

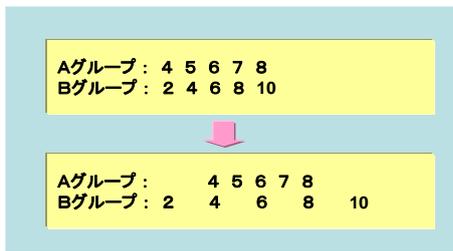
『そこから、何を感じとるか』を重視していきますので、表題にありますように、四則計算とルート of 計算が出来る程度の電卓を準備してお読みいただければ用は足ります。私自身、某大学(名前を出せば皆さんご存知だと思いますが・・・)で、記述式の試験において 40 文字程度の解答で統計学の単位を取っています。おそらく、長い歴史の中なかで最小解答文字数に近いのではないのでしょうか？
ただただ担当された先生の懐の広さに感服する次第ではありますが・・・>

ようするに、統計学のみならず『いかに的を得るか』が重要なのでしょう。本連載は、表題にもありますように、数学的な厳密さよりも直感的に理解することに重点を置いた表現を用いています。ご専門でやられている先生がたには、不満もあろうかと思いますが、本連載の主旨をご理解いただければ幸いです。

おおまかに！ - I -

直感的に理解？

下に 5 個の数字からなる、2 つのグループがあります。皆さん、2 つのグループを比べると、どちらがどのくらいバラツキが大きいと思いますか？並べ変えると・・・多くの方が『B は A より 2 倍ぐらいばらついているかなあ』と感じるのではないのでしょうか。バラツキを計算すれば、その通りなのですが・・・直感に客観性を持たせるのが、統計だと思ってください。いまは、『2 倍ぐらいばらついているかなあ』という感覚を大切にしてください。この解法は、連載の中で説明します。実は、この程度であれば紙と鉛筆で解けるのです。電卓はいりません。



お友達、集まれ！

ここでもう 1 つ、直感のお話をしましょう。今度は、一桁の数字がいっぱい並んでいます。これから何かを感じますか？私は・・・めまいがするだけで、何も感じませんねえ。では、先ほどと同じように並び変えてみましょうか。うーん。なんだか、5 が多そうですね。

3 6 4 5 5 6 7 6 5 8 6 5 7 6 5 6 4 5 6
6 4 9 5 5 1 7 4 5 2 2 4 5 5 6 4 7 5
3 4 8 5 4 5 5 6 5 7 5 5 7 6 3 4 6 7 9
3 4 7 6 6 5 7 4 4 8 4 5 6 3 5 5 4 4
1 7 5 6 6 5 7 4 6 5 5 8 3 3 7 5 7 7 6
4 7 6 5 4 4 5



1 1 2 2 3 3 3 3 3 3 4 4 4 4 4 4 4 4
4 4 4 4 4 4 4 4 4 4 4 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
6 6 6 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 8
8 8 8 9 9

では、個々の数字の数を数えて表にしてみましょう。確かに、5 が一番多い事が確認できます。

値	度数	累積度数
1	2	2
2	2	4
3	7	11
4	19	30
5	29	59
6	20	79
7	15	94
8	4	98
9	2	100

このように、生データから、何かその分布の特徴や癖を引き出すための手法が『統計』という手法なのです。

統計で行われるのは、『縮約』と呼ばれる方法です。縮約とは、「データとして並んでいるたくさんの数字を、何かの基準で整理整頓して、意味のある情報だけを抽出する」ということを意味する言葉で、おおまかに次の 2 つの手法があります。

① グラフ化してその特徴をとらえる。

② 1 つの数字で特徴を代表させる。

その代表する数字のことを『統計量』と呼びます。

先ほどのように、同じ数値のデータが何個あるか数えて表にまとめたものを 度数分布表(Frequency Table)といいます。

ここで、データの個数を『度数』といい、度数を足していったものを『累積度数』といいます。

身長、体重のような場合、同じ数値が複数回出てくることは少ないので、適当な小区間に分割し、各小区間に何個のデータがあるかを数えて表にすることがあります。このときの各小区間を『階級』と言います。

ヒストグラム

