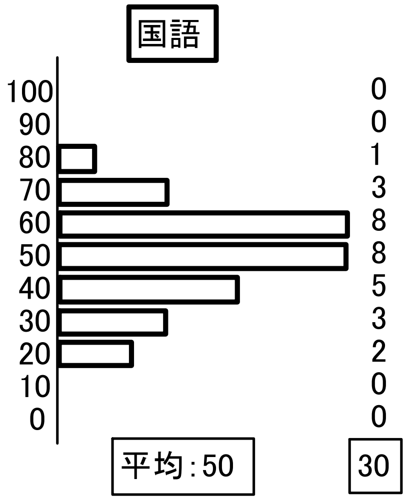
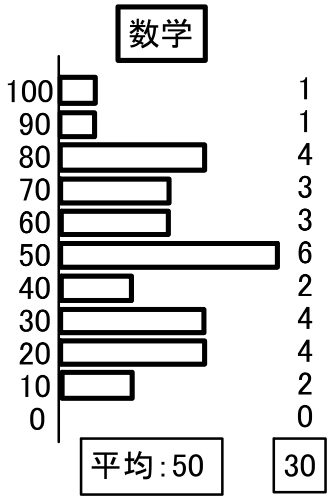
§16. 語る前に知っておきたい（偏差値の意味）

偏差値の意味について考えてみましょう。まず、具体的な例を取り上げます。ある生徒（A 君とします）が国語と数学のテストを受けたとします。どちらも 100 点満点のテストで、この生徒は国語も数学も 70 点でした。平均点はどちらもちょうど 50 点です。果たして A 君は、国語も数学も同じ程度の出来（全体から見て）だと言えるでしょうか。

これだけのデータでは、誰でも首を傾げたくなるでしょう。まず、「同じ程度の出来」という言い方がかなりあいまいなので、もう少し分かりやすい表現の仕方で言い替えておきます。「全体のどのあたりに位置する成績か」という意味であるとします。学生時代に何度もテストを経験してきた人なら誰でも同じだと思いますが、それぞれのテストにおける得点の分布がどのようになっているかということが気になるところです。それでは、国語と数学の得点分布が図のようだったとします。ただし話を単純化するために得点は 10 点きざみにしています。





この二つのグラフを見せられれば、誰でも A 君は国語の方が数学よりも出来が良かったと考えるでしょう。ことばで表現すれば、「国語の方が数学より得点のばらつきが小さいので、平均点、得点が同じでも、国語の方が良い成績だったと言える」ということになります。このように、得点分布を表すグラフ（あるいはそのもととなる表）があれば、テストにおける出来具合は一目瞭然で分かりやすいです。しかし、テストを何回も繰り返し、データとして積み重ねていかなければならないような場合には、いちいち得点分布表やグラフを添付するというのは、あまり合理的と言えません。

今は、一つの教科の得点に関する話でしたが、例えば 5 教科の合計で出来具合を評価する場合にはどうでしょう。それぞれの教科の素点を合計するのではあまりにも乱暴ですし、誰しも疑問や不満を抱くのではないでしょうか。

いずれの場合も、もしも各教科の得点を適正に客観的な数値に変換できれば、大変好都合です。そこで登場するのが偏差値です。勿論、得点分布表のような的確な資料とは言えませんが、たった一つの数値で、ある程度のことを読み取ることが可能です。言うまで

もありませんが、そのためには偏差値の意味をしっかり理解しておく必要があります。
それでは偏差値はどのように求められるのかを見ていきます。上の説明にもあったように、偏差値は得点分布のばらつき具合が問題になりますが、そのばらつき具合を数値化する計算法を上の国語の得点分布を例に示します。
まず平均点は 50 点ですが、一つひとつのデータ（得点）がこの平均点からどれだけ隔たっているか（即ち平均点との差）を計算して、足し合わせていきます。得点が平均点より小さいときと大きいときとでは、引き算を逆にしなければならないので、（得点－平均点）^{2} というようにして、足す値が常に＋になるようにします。国語の場合で実際に計算するときの仕方を示します。全員分の平均からの隔たりを二乗して合計します。

$(20-50)^2\times2+(30-50)^2\times3+(40-50)^2\times5+(50-50)^2\times8+(60-50)^2\times8$
$+(70-50)^2\times3+(80-50)^2\times1$
$=1800+1200+500+0+800+1200+900$
$=6400$

この合計値は人数に関係しているので、人数で割ります。

$6400/30\fallingdotseq213.3$

平均との隔たりをそれぞれ二乗しているので、（関数電卓を用いて）平方根を求めます。

$\sqrt{213.3}\ \fallingdotseq14.6$

こうして求められた数値を「**標準偏差**」と言います。記号は σ（ギリシャ文字でシグマと読む）を用いるのが一般的です。
さていよいよ偏差値ですが、今求めた標準偏差を用いて得点 70 に対する偏差値は次のように計算されます。

$(70-50)/(\sigma/10)+50$　　注：()の中の 50 は平均点の 50、最後の 50 は偏差値計算の公式としての 50
$\fallingdotseq63.7$

偏差値を計算する一般的な公式は、次のようになります。y は求める偏差値、x は得点、m は平均点、σ は標準偏差です。

$\mathbf{y=(x-m)/(\sigma/10)+50}$ ・・・ ①　　注：$y=10(x-m)/\sigma+50$ と考えてよい。σ/10 とする意味については後程

それでは上の数学で同じように計算してみましょう。

$(10-50)^2\times2+(20-50)^2\times4+(30-50)^2\times4+(40-50)^2\times2+(50-50)^2\times6+$
$(60-50)^2\times3+(70-50)^2\times3+(80-50)^2\times4+(90-50)^2\times1+(100-50)^2\times1$
$=3200+3600+1600+200+0+300+1200+3600+1600+2500$
$=17800$
$17800/30\fallingdotseq593.3$
$\sigma=\sqrt{593.3}\fallingdotseq24.4$

これから、数学の得点 70 に対する偏差値 y は

$y=10(70-50)/24.4+50$
$\fallingdotseq58.2$

国語、数学のそれぞれの偏差値を実際に計算して比較すると、偏差値の意味合いが実感できると思います。ばらつきの度合いを数値化した σ の値によって得点（－平均）を割る

ので、同じ得点 70 に対しても、σ が小さければ偏差値は大きくなるし、逆に大きければ偏差値は小さくなります。

後回しになりましたが、偏差値を求める式①について少し説明を加えます。まず式の最後にある＋50 ですが、(もうお気づきと思いますが) 平均点の偏差値を 50 にするためです。得点が平均を下回れば、偏差値は 50 より小さくなります。次に割り算の分母ですが、σ そのものでなく σ/10 としています。これは結果的に分数の値を 10 倍することになりますが、偏差値の幅を拡大する働きをしています。例えば、偏差値が 65 の場合(15＋50)、この ×10 が無いと 1.5＋50 の 51.5 なってしまいます。つまり、虫眼鏡のような役目と言えます。

偏差値に関することで二、三補足しましょう。高校・大学受験のことが話題になるとき、この偏差値ということばを多く耳にするようになります。偏差値絶対主義だとか、偏差値の一人歩きというような、良くない言い方も聞かれます。確かに偏差値の本当の意味を知らないで、その数値にばかり振り回されるのはどうかと思います。ひどいときには、まるで柔道選手が獲得した段位のように、「○○君は偏差値 65 の能力を有する」というような言い方をする人がいます。当たり前ですが、同じ実力を発揮できたとしても、テストを受けたときの集団の能力が高いか低いかで偏差値は異なります。出来の良いメンバーの中で受ければ、普段学校で 65 位の偏差値が取れていたとしても、50 を割ってしまうかもしれません。柔道選手の二段はどこへ行っても二段です(勿論全国共通のレベルでの基準があっての話ですが)。

もう一つ知っておきたいことは、偏差値を扱う場合には、ある程度のデータ量が欲しいということです。つまり、あまり人数の多くないところでは、偏差値が本来の意味での適切な値を示さなくなります。例えば一クラス(40～50 人)程度の場合には、得点分布がテストのたびに不規則なでこぼこ状態になりがちですが(上の数学がまさしくその例)、本人の能力は変わらなくても、偏差値がころころ変動する可能性があります。学年全体(数百人)程度になると、得点分布が小高い丘のような状態に近づきます(ばらつきが大きいときは裾野が広がるなだらかな山、ばらつきが小さいときはヘアピンのような急な山)。こうした中でならば、偏差値はかなり意味をもってくるでしょう。そして、理想的には、平均点を中心として左右対称に山のように分布する状態です。県や国のレベルで実施されるテストで、出題が適切であればそのような状態になるはずです(左右非対称の山になったり歪んだ山になったりするのは、出題の仕方が良くない！のです)。この場合には、偏差値だけで自分が全体のおよそどれくらいの順位にいるのかも分かります。本来、偏差値はそうした使われ方をするべきものなのでしょうが、現実にはかなりの小集団でも活用され、時として我がもの顔で歩き出したりもするようです。

最後に、関心を持った人のために、実際的なところでの標準偏差 σ の求め方を紹介します。まず統計処理の機能を持った関数電卓があれば、全員の得点を打ち込むことで平均、標準偏差を求めることができるので、偏差値も簡単に求められます。

そうした関数電卓がない場合は普通の電卓を用いて計算することになりますが、次のような工夫(σ を求める式の変形)をすることで、いくらか楽に済ませることができます。

次のような度数分布表があったとします。

変数　x_1　x_2　x_3　・・・　x_i　・・・　x_k
度数　f_1　f_2　f_3　・・・　f_i　・・・　f_k　　合計 $N=f_1+f_2+f_3+\cdots+f_i+\cdots+f_k$
平均　$m=(f_1x_1+f_2x_2+\cdots f_kx_k)/N$

$\sigma^2=\{f_1(x_1-m)^2+f_2(x_2-m)^2+f_3(x_3-m)^2+\cdots+f_i(x_i-m)^2+\cdots+f_k(x_k-m)^2\}/N$
$=(f_1x_1^2+f_2x_2^2+\cdots+f_kx_k^2)/N-2m(f_1x_1+f_2x_2+\cdots f_kx_k)/N+m^2(f_1+f_2+\cdots+f_k)/N$

$(f_1x_1+f_2x_2+\cdots f_kx_k)/N=m$、$(f_1+f_2+\cdots+f_k)/N=1$　だから
$\sigma^2=(f_1x_1^2+f_2x_2^2+\cdots+f_kx_k^2)/N-2m^2+m^2$
$=(f_1x_1^2+f_2x_2^2+\cdots+f_kx_k^2)/N-m^2$　・・・　②

このことから、度数分布表に手を加えて、

変数	x_1	x_2	x_3	・・・	x_i	・・・	x_k
	x_1^2	x_2^2	x_3^2	・・・	x_i^2	・・・	x_k^2
度数	f_1	f_2	f_3	・・・	f_i	・・・	f_k
	$f_1x_1^2$	$f_2x_2^2$	$f_3x_3^2$	・・・	$f_ix_i^2$	・・・	$f_kx_k^2$

のような表を作れば、②から σ^2 の値をいくらか楽に計算することができます。