必修「数学Ⅰ」で統計を学ばせるための教材実践研究

一橋大学 経済学研究科　駒野誠
makomano@mbj.nifty.com

１．研究の動機

筆者は、「数学は、無限を掴む」が大テーマであると考え、そのためには、「無限を有限化する行為が欠かせず、また、逆に有限個を無限にするという、他に類を見ない学ぶべき価値がある」と考え生徒・学生に指導してきた。

無限のなかに‘共通点’を見いだし、それによって、無限から有限の世界に誘うことが可能となる。このことは人間の思考効率化の特徴といってよいだろう。これは、AI そのものの考え方かもしれない。

しかし、現在の学校教育での統計指導は、多くの定義・公式を適用するいわば暗記ものになりがちである。（統計に関わらず、答を速く出す算術が主眼になっているように感じる。ここで、塩野直道の要目カス論：「要目として羅列してある数学的な話題は，生徒が数学的な活動をして後に残った食べカスに過ぎない。活動の質自体が重要な教育的内容である。」を思い出す。）答えを出す、答は一つが数学と思い込まされてはいないだろうか。

「見えなかったモノが観えるようになる」を狙った、モノの見方に新たな視点をプラスする数学としての力を生徒につけさせられるとの認識が持てないためか、現場での統計授業に力が入っていない現状がある。入試でも、2022 年度以降の高大連携で、すぐにでも実行可能で効果あることとして、各大学の入試において、受験生の自分のノート１冊持ち込み OK の決断で、入試は新たな歩を踏み出せるだろう。

偏差の平方の平均、2 乗の平均から平均の 2 乗を引いたもの、が分散だ！覚えて置けとヤクザ調になったりする。暗記はある意味、考えることを放棄し、考えないで済むという利点があるともいえるが、数学科の存在意義は、算数の「…算」のような答えを出すマニュアル的算術方法の暗記ではない。これを問題解決だとの勘違いがあるのかもしれない。公式の文字に数を代入することが数学であると生徒に思わせていたらダメだろう。

２．学校数学の役割

{‘かぞえる’，‘くらべる’，‘はかる’，‘かえる’} を四面体構造の頂点とする学校での数学教授の中心的な概念が学校教育での dependable(頼りになる)とする。その内部で、困ったときの対処法に、portable(持ち運び可能)な {仲間をつくる，次元をあげる，相棒をみつける，定義をする} の実行を試みる。また、これら全体を通じ、個々の数式・定理・公式らを従え、最上位にあるべきは、‘共通点’を見出すことが sustainable(持続可能)と考える。数学教育という塀の中での村社会にならないような生き生きとした学びが肝要である。

３．問題点

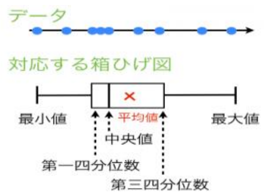
新学習指導要領では、‘データの分析’が引き続き「数学Ⅰ」で、「数学 B」では、ベクトルを「数学 C」へ追い出し、‘数列’，‘統計的な推測’，‘数学と社会生活’が入る。必修の「数学Ⅰ」で統計を扱う際に、未習の’ベクトル’、’数列’の使用禁止の影響が大である。具体的には、添え字と Σ である。

＜箱ひげ図の扱い＞

2022 年からは、‘四分位数’や‘箱ひげ図’が中学校へ移行となる。しかし、「数学 A」の‘数学と人間の活動’の章の中の整数の性質の学習で、もっと数学的に扱うべきことがある。

○「数学 A」整数とリンクさせる指導例

中学校へ移行した理由は、簡単なことだから中学へ、が理由だとするならば、それは誤りである。



その理由を示そう。四分位数と箱ひげ図（box-plot

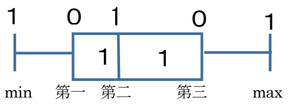
diagram) は、データを小さい値の順にソーティングし、4 等分に分割し、その位置にくる数を「四分位数」といい、小さい方から「第 1 四分位数Q_1」、「第 2 四分位数Q_2」(中央値)、「第 3 四分位数Q_3」と呼ぶ。

Q は Quartile （4 分の 1 は quarter）の頭文字で、四分位数(quartile) first-, second-, third- ただし、4 等分の位置にデータがない場合もあり、その時は両隣のデータの平均値を四分位数とするなど、見た目ほど簡単ではない。箱にはデータが均等に分布しているわけではないことは特にわかりにくい。

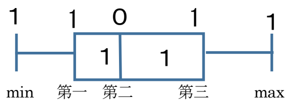
指導例　データの個数に関して、半分半分とするのだから、mod 4 つまり 4 で割った余りが 0，1，2，3 の 4 通りを考えさせることで、すべてを掴んだことになる。データの個数をn としたときの図の記号は、｜が両隣の平均を四分位数とするとき、●はデータそのもの

$n \equiv 0(mod\ 4)$　○ ｜ ○ ｜ ○ ｜ ○

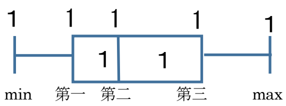
$n \equiv 1(mod\ 4)$　○ ｜ ○ ● ○ ｜ ○



$n \equiv 2(mod\ 4)$　○ ● ○ ｜ ○ ● ○



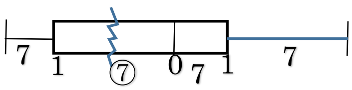
$n \equiv 3(mod\ 4)$　○ ● ○ ● ○ ● ○



よくある問題

問題　データの数と箱ひげ図が与えられたとき、4 等分のところにデータが存在しているか否かが重要であり、また、Q_2が箱の中を仕切っているが、Q_1とQ_2の間にカーテンℓを入れたとき、ℓよりも小さい数がどのように分布しているかは不問である。そのため、ℓより小さいデータの最小値と最大値という問題が生じる。　そこで、すべてを把握するために、データの個数で分類するのが数学である。

30 人のデータの箱ひげ図が下図である。



30 個の分布　30≡2(mod 4) カーテンℓより小さいデータの個数 N は、8≦N≦15　なぜなら、⑦の 7 個が、箱内でℓより左に全部あるときが最大個数で 15 個、⑦の 7 個全部が箱内ℓの右かつQ_2の左にあるときが最小個数の 8 である。

４．現状認識

(1)　教科書のあり方

繰り返すが、必修の「数学Ｉ」では、新「数学 C」の‘ベクトル’ で学習するベクトルの大きさや内積、新「数学 B」で学ぶ‘数列’での Σ 記号などを用いることはできないことを前提に教材を構成する必要がある。そのため、「数学Ｉ」での、‘データの分析’における式変形は、数学が得意な生徒たちには Σ 記号を導入して済ませられるが、数学が苦手な生徒たちには Σ 記号を短時間で差し挟むことは指導上無理があり、Σ は使えないために、平均のたびに、$\frac{x_1+x_2+\cdots+x_n}{n}$ のように（ +⋯+ ）/ n を多用しなければならず先生方は指導に苦労している現状がある。

筆者がこの打開策を模索していて、共通点の‘平均’に気が付いたとき、ワクワク感は最高潮であった。やはり、‘共通点’は思考の根幹に位置づけられる。

(2)　定義と式変形

統計の代表値である、‘最頻値’、‘中央値’などは、データの和とは無関係であり、小中学校で済む内容である。その後の代表値としての「数学Ｉ」での内容は、‘分散’、‘標準偏差’、‘共分散’、‘相関係数’だが、、‘回帰直線の傾き(回帰係数)’なども学ばせたい。ここで暗記モノではない、数学としての統計をしたいものである。

「数学Ｉ」での‘データの分析’における共通点は、平均を利用した値の処理であることである。

この‘共通点’を見出すことこそ、数学学習の最大の課題ともいえる。純粋数学においても、共通点から同値関係を考え、分類作業が始まる。この論文の肝は、いかにして、Σ を使わずに指導できるのか模索していたところ、平均を計算する式たちに‘共通点’があり、Σ を使わずに種々の公式を変形することが容易にできたので、高校生、大学生、教員に実践してみた結果、好評であったので報告する。この論考では、念のため Σ など一部両方併記しておく。

５．表現方法

(1)　データの表現

数学として、最初に困るのが、資料、データxという表現である。集合なのか、順序対なのか、不明であること。集合ならば、データ$x=\{x_i\}$ $(i=1,2,3,\ ・・・,n)$,$y=\{y_i\}$ $(i=1,2,3,\ ・・・,n)$ とするのだろう。集合よりも、順序対のほうがよいのではないか。

順序対$x=(x_1,x_2,\ ・・・,x_n),y=(y_1,y_2,\ ・・・,y_n)$

といっても、x_iという、「数学B」の数列で学習する添え字付きの数(関数)の前に使用しているという、

そこで、ベクトルに似せて、$x=(x_1,x_2,\ ・・・,x_n)$などと書き、xのメンバーたちと呼ぶことにする。

異なる種類のデータを扱うときやそこから組を作るには順序が肝心となる。また、後に学ぶ、ベクトル、数列につながるものでなければならない。リンクで見出す **'共通点'** を有する知識でなければ、portable なモノにならず、社会に出て役に立たない。

(2)　平均をとる操作 bar

ここでは、これを‘平均化’と呼ぶことにする。

Bar は、否定として**ド・モルガンの法則**$\overline{A\cap B}=\overline{A}\cup\overline{B}$は、bar の配下にある 3 つを個々に否定したような、使い勝手がよい法則である。次のような括弧の代用としての使い方もある。例）$x^2=(\overline{x-p}+p)^2=(x-p)^2+2p(x-p)+p^2$　のようにテーラー展開では活躍する。

平均化では、bar を使うことの凄さは、添え字を使うことなく、Σ も使う必要がないことである。

データxの平均 $\overline{x}=\frac{x_1+x_2+\cdots+x_n}{n}$　・・・　$\frac{1}{n}\sum_{i=1}^{n}x_i$を 1 文字$\overline{x}$で表現できることである。凄い！

ことばに直せば、この「xの平均」とは「xのメンバーたちの平均をとる」ことをいう。平均値とは、和をとることが前提で、足した個数で割って得られる数のこと。

平均化の公式（bar）線形である：

① $\overline{ax+by+c}=a\overline{x}+b\overline{y}+\mathrm{c}$　,

② $x=y$ならば$\overline{x}=\overline{y}$　,　③ $\overline{(\overline{x})}=\overline{x}$

この等式は後に度々使用するので、ここで 1 度だけ$\frac{1}{n}(\ldots+\cdots+\cdots)$を用いて証明しておく：

証明）①　通常の証明は、Σを利用して、

左辺＝$\overline{ax+by+c}=\frac{1}{n}\sum_{i=1}^{n}(ax_i+b\,y_i+c)=$

$a\frac{1}{n}\sum_{i=1}^{n}x_i+b\frac{1}{n}\sum_{i=1}^{n}y_i+\frac{1}{n}\sum_{i=1}^{n}c=a\overline{x}+b\overline{y}+\mathrm{c}=$ 右辺

◎シグマを使わない指導による証明

$\frac{1}{n}\{(ax_1+by_1+c)+(ax_2+by_2+c)+\cdots+(ax_n+by_n+c)\}$

$=a\times\frac{x_1+x_2+\cdots+x_n}{n}+b\times\frac{y_1+y_2+\cdots+y_n}{n}+\frac{c+c+\cdots+c}{n}$

$=a\overline{x}+b\overline{y}+\mathrm{c}$　　//

②　$x=y$より、任意のiで$x_i=y_i$であるから、平均 $\overline{x}$と $\overline{y}$ は等しい。

③　平均すると、1 つの定数になるので、平均しても変化なし。//

よく知られている確率変数の変換である公式：

期待値　$\mathrm{E}(ax+b)=aE(x)+b$　　,

分散 $V(ax)=a^2V(x)$

この証明は、期待値 $\mathrm{E}(ax+b)=\overline{ax+b}=a\overline{x}+b$,

xの分散が$\mathrm{V}(x)=\overline{(x-\overline{x})^2}$ であるから、

axの分散 $\mathrm{V}(ax)=\overline{(ax-\overline{ax})^2}=\overline{\{a(x-\overline{x})\}^2}$

$=a^2\overline{(x-\overline{x})^2}=a^2\mathrm{V}(x)$　//

などにも容易に対応できる。

また、教科書(「数研出版」研究‘変量の変換’)で、変量xから変量yへの変数変換:$y=ax+b$の平均と分散の計算を紹介している。

ここでは、平均化を用いて計算してみる。

平均は、$\overline{y}=\overline{ax+b}=a\overline{x}+b$

分散は、$\sigma_y{}^2=\overline{(y-\overline{y})^2}=\overline{\left((ax+b)-(a\overline{x}+b)\right)^2}$

$=\overline{\{a(x-\overline{x})\}^2}=\overline{a^2(x-\overline{x})^2}=a^2\overline{(x-\overline{x})^2}=a^2\sigma_x{}^2$

1 ページを割いて掲載されたことが、たったの数行で計算終了となるほど容易になった。

主たる統計量の表つくり

データ	x		y		$x-\overline{x}$		y−$\overline{y}$		$(x-\overline{x})(y-\overline{y})$		$(x-\overline{x})^2$		$(y-\overline{x})^2$		xy		x^2		y^2		$\frac{x-\overline{x}}{\sigma_x}$	$\frac{y-\overline{y}}{\sigma_y}$	$\frac{x-\overline{x}}{\sigma_x}\times\frac{y-\overline{y}}{\sigma_y}$
	メンバー		メンバー		xの偏差		yの偏差		偏差の積		(xの偏差)の平方		(yの偏差)の平方		積		平方		平方				
A	x_1	1	y_1	4	$x_1-\overline{x}$	−2	$y_1-\overline{y}$	−1	$(x_1-\overline{x})(y_1-\overline{y})$	2	$(x_1-\overline{x})^2$	4	$(y_1-\overline{y})^2$	1	x_1y_1	4	$x_1{}^2$	1	$y_1{}^2$	16	$-\sqrt{2}$	$-\sqrt{5}/\sqrt{6}$	$\sqrt{5}/\sqrt{3}$
B	x_2	2	y_2	5	$x_2-\overline{x}$	−1	$y_2-\overline{y}$	0	$(x_2-\overline{x})(y_2-\overline{y})$	0	$(x_2-\overline{x})^2$	1	$(y_2-\overline{y})^2$	0	x_2y_2	10	$x_2{}^2$	4	$y_2{}^2$	25	$-1/\sqrt{2}$	0	0
C	x_3	3	y_3	5	$x_3-\overline{x}$	0	$y_3-\overline{y}$	0	$(x_3-\overline{x})(y_3-\overline{y})$	0	$(x_3-\overline{x})^2$	0	$(y_3-\overline{y})^2$	0	x_3y_3	15	$x_3{}^2$	9	$y_3{}^2$	25	0	0	0
D	x_4	4	y_4	4	$x_4-\overline{x}$	1	$y_4-\overline{y}$	−1	$(x_4-\overline{x})(y_4-\overline{y})$	−1	$(x_4-\overline{x})^2$	1	$(y_4-\overline{y})^2$	1	x_4y_4	16	$x_4{}^2$	16	$y_4{}^2$	16	$1/\sqrt{2}$	$-\sqrt{5}/\sqrt{6}$	$-\sqrt{5}/2\sqrt{3}$
E	x_5	5	y_5	7	$x_5-\overline{x}$	2	$y_5-\overline{y}$	2	$(x_5-\overline{x})(y_5-\overline{y})$	4	$(x_5-\overline{x})^2$	4	$(y_5-\overline{y})^2$	4	x_5y_5	35	$x_5{}^2$	25	$y_5{}^2$	49	$\sqrt{2}$	$2\sqrt{5}/\sqrt{6}$	$2\sqrt{5}/\sqrt{3}$
データ数	5		5		5		5		データ数	5	データ数	5	データ数	5	データ数	5	データ数	5	データ数	5	5	5	5
和	15		25		0		0		和	5	和	10	和	6	和	80	和	55	和	131	0	0	$5\sqrt{5}/2\sqrt{3}$
平均	$\overline{x}$	3	$\overline{y}$	5	$\overline{x-\overline{x}}$	0	$\overline{y-\overline{y}}$	0	$\overline{(x-\overline{x})(y-\overline{y})}$	1	$\overline{(x-\overline{x})^2}$	2	$\overline{(y-\overline{y})^2}$	1.2	$\overline{xy}$	16	$\overline{x^2}$	11	$\overline{y^2}$	26.2	0	0	$\sqrt{5}/2\sqrt{3}$
分散									σ_{xy}		σ_x^2		σ_y^2		σ_{xy} $=\overline{xy}-\overline{x}\cdot\overline{y}$		σ_x^2 $=\overline{x^2}-\overline{x}^2$		σ_y^2 $=\overline{y^2}-\overline{y}^2$				$r_{xy}=\sigma_{xy}/\sigma_x\sigma_y$
標準偏差	1.41		1.10							1	σ_x	$\sqrt{2}$	σ_y	$\sqrt{1.2}$									相関係数

$\overline{x}=3$　,　$\overline{y}=5$　,　$\overline{(x-\overline{x})^2}=s_x{}^2=\sigma_x{}^2$　,　$\overline{(y-\overline{y})^2}=s_y{}^2=\sigma_y{}^2$,　$\overline{(x-\overline{x})(y-\overline{y})}=s_{xy}=\sigma_{xy}$ など。

$\boxed{\text{定義}}$　　xの分散とは、偏差の平方の平均

つまり、$s_x{}^2=\sigma_x{}^2=\overline{(x-\overline{x})^2}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2$

xの標準偏差とは、xの分散の平方根である。

$$s_x = \sigma_x = \sqrt{\overline{(x-\overline{x})^2}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2}$$

x, yの共分散とは、x, yの偏差の積の平均である。

$$s_{xy} = \sigma_{xy} = \overline{(x-\overline{x})(y-\overline{y})} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})$$

(3)　記号の表現方法

x, yの共分散s_{xy} のディメンジョンがわかるには、x, xの共分散がxの分散と言い換えられることである。

そこで、$s_{xx} = s_{x^2} = {s_x}^2 = {\sigma_x}^2$ 等と表記できるが、標準偏差$s_x = \sigma_x = \sqrt{s_{xx}} = \sqrt{s_{x^2}} = \sqrt{S_x^2}$ 等と複数の表現ができるが、どれも一長一短である。

(4)　平均の式変形

本稿では、言葉の代わりにΣ計算も書いたが、「数学Ⅰ」の授業ではΣは使わずに、定義の言葉通りの計算を bar だけの計算で処理できるので、これまでになく画期的であると自負している。

①xの偏差$x-\overline{x}$の平均＝$\overline{x-\overline{x}}$＝$\overline{x-\overline{x}} = \overline{x} - \overline{x} = 0$

②xの分散＝偏差の平方$(x-\overline{x})^2$の平均＝$\overline{(x-\overline{x})^2}$

$$\begin{aligned}{s_x}^2 &= \overline{(x-\overline{x})^2} = \overline{x^2 - 2x\overline{x} + (\overline{x})^2} \\ &= \overline{x^2} - 2\overline{x}\cdot\overline{x} + (\overline{x})^2 = \overline{x^2} - (\overline{x})^2\end{aligned}$$

この変形は、数学が苦手な生徒にも、得意な生徒にも、先生方にも好評である。

2 変数変換$z = x + y$の分散も容易に計算できる；

$$\begin{aligned}{s_z}^2 &= \overline{\left\{(x+y) - \overline{(x+y)}\right\}^2} = \overline{\{(x-\overline{x}) - (y-\overline{y})\}^2} \\ &= \overline{\left\{(x-\overline{x})^2 - 2\overline{(x-\overline{x})(y-\overline{y})} + (y-\overline{y})^2\right\}^2} \\ &= \overline{(x-\overline{x})^2} - 2\overline{(x-\overline{x})(y-\overline{y})} + \overline{(y-\overline{y})^2} \\ &= {s_x}^2 - 2s_{xy} + {s_y}^2\end{aligned}$$

次に、相関係数や回帰直線を学ぶには、標準化の考えが欠かせない。しかし、教科書では省かれ、相関係数が暗記すべき公式となっている。しかし、bar をとる計算によって、変換式などのチェックが容易になった。

(5)　「覧古考新」

「覧古」とは、Met Before(既習)のことであるが、その後の成長によって、振り返ってみると、その当時は気が付かなかったことが深く観えて、新たなことを考えだせる「考新」となること。

例）中学での直線 $y = ax + b$…①の学習が 2 点通過型で、連立方程式による係数決定に偏っていることである。それは、パラメータの消去方法として、通る点(1, 3)があれば、$3 = a \times 1 + b$…②

ここで、①−②で、パラメータの１つのbを消去し、

$y = ax + (3-a) = a(x-1) + 3$　 つまり、

通る点(1，3)の座標を‘見える化’した式： $y - 3 = a(x-1)$ を紹介する指導がなされていないことに気付かされる。後の接線の公式の指導で苦労することになる。

x 軸、y 軸上の両方に並べたデータをどうやって組にするかという根本の問題である。

従来の「xが 1 のときyが 3、xが 4 のときyが 9 である 1 次関数？」だけでなく、新種の問い「ある日、新芽 A の長さは1㎝で 1 週間後は 4 ㎝、新芽 B の長さは 3 ㎝で 1 週間後は 9 ㎝であった。」

データをどのように組にして、2 次元座標平面上にプロットするのか？

ここで、これらの芽の伸びが 1 次関数的だとすると、2 種類の新芽 A, B の長さを組にし、点(x, y)と見做したとき、ある日の長さは点(1，3)で、その 1 週間後は点(4，9)と見做して整理できる。

$y - 3 = \frac{6}{3}(x-1)$　⇔　$\frac{y-3}{6} = \frac{x-1}{3}$　⇔　$y - 3 = 6t, x - 1 = 3t$(tは実数)　と表現できる。

(6)　標準化

‘標準化’とは、データを平均＝0，標準偏差＝１(分散＝1)に変換することである。数学的には‘正規化’と呼ぶのがよいだろう。異なる２種類のデータを扱うには、単位の統一が欠かせない。つまり、正規直交座標系に点をプロットするためには、単位を統一する必要がある。

例）高校生の身長と体重では大方 3 桁、2 桁と異なるので、模式図となる見える化作業では、標準化が必要となる。エクセルでは、変換処理は容易である。

変換　$x = (x_1, x_2, ・・・, x_n)$　→

z= $(z_1, z_2, ・・・, z_n)$ の定義を$z = \frac{x-\overline{x}}{\sigma_x}$とする。

xの平均が$\overline{x}$，標準偏差がσ_xであるデータを、zの平均が 0，標準偏差が 1（分散も 1）のデータにxからzへ線形変換したものである。　‘見える化’する。

これは、データの組P(x, z),2 点 A$(\overline{x}, 0)$, B$(\overline{x}+\sigma_x, 1)$を通過する直線であり、その式は、

$x = \overline{x} + \sigma_x t$ ，$z = 0 + t$（tは実数）であり、$z - 0 =$

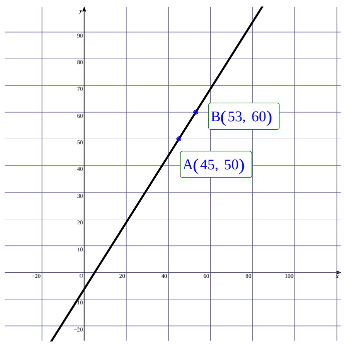
$\frac{1}{\sigma_x}(x-\bar{x})$ である。変換式は、$\mathbf{z}=\frac{x-\bar{x}}{\sigma_x}$ は標準測度とも呼ばれている。

$$\text{zの平均}=\bar{z}=\overline{\left(\frac{x-\bar{x}}{\sigma_x}\right)}=\frac{\bar{x}-\bar{x}}{\sigma_x}=0$$
$$\text{z の分散}=\overline{(z-\bar{z})^2}=\overline{(z-0)^2}$$
$$=\overline{z^2}=\overline{\left(\frac{x-\bar{x}}{\sigma_x}\right)^2}=\frac{\overline{(x-\bar{x})^2}}{{\sigma_x}^2}=\frac{{\sigma_x}^2}{{\sigma_x}^2}=1$$

(7)　偏差値（標準化の一例）

Web 上でも、‘偏差値’についての公式を紹介しているが、変換の立場での記述はないようである。
ここでのキーワードは、‘標準化’と‘線形変換(一次変換)’である。数学科の指導としては、生徒自身が自分自身で公式が作れるような説明を有していることが重要である。

標準化による変数変換の具体的な使用例が‘偏差値’である。例）xの平均 45 点、標準偏差 8 点のとき、これに対応するyの平均を 50、標準偏差を 10 とする線形変換が偏差値である。 $x=45+8t\,,y=50+10t$ であり、対応するデータを組とし、平面上の点と見做せば、これは 2 点$(45{,}50),(45+8{,}50+10)$を通る直線を求めることになる。$y-50=\frac{10}{8}(x-45)$ つまり、 $y=50+10\times\frac{x-45}{8}$ である。



$z=\frac{x-45}{8}$で、標準化していることを確認してみよう：

$\text{zの平均}\ \bar{z}=\overline{\left(\frac{x-45}{8}\right)}=0,$

$z\text{の分散}=\overline{(z-\bar{z})^2}=\overline{\left(\frac{x-45}{8}-0\right)^2}=\overline{\left(\frac{x-45}{8}\right)^2}=\frac{8^2}{8^2}=1$

∴ zの標準偏差＝1

ここで、$z=\frac{x-45}{8}$は、xの平均 45 点、標準偏差 8 点を z の平均 0、z の標準偏差 1 への線形変換である。

なぜなら、2 点$(45{,}0),(45+8{,}0+1)$ を通る直線 $z-0=\frac{1}{8}(x-45)$ である。すなわち、$z=\frac{x-45}{8}$

(8)　相関係数

相関係数$r=\frac{\sigma_{xy}}{\sigma_x\sigma_y}$ この定義はどのように見出したのか？ ベクトルの内積の考えに依らず考えてみよう：

＜ア＞ $z=\frac{x-\bar{x}}{\sigma_x},w=\frac{y-\bar{y}}{\sigma_y}$ とおくとき、$w=rz$の係数rの意味を考える

従来の計算は、誤差の平方の和、${\epsilon_i}^2=\sum_{i=1}^{n}(w_i-rz_i)^2$の最小値を平方完成（または微分）で求めることでした。この回帰変数r（回帰直線の傾き）を‘最小２乗法’を利用して求めてみよう：しかし、そこでの計算にはΣがあり面倒になるので、平均の最小値を考えることにする。ここが Point である。
＜数学の問題解決は、与えられたままや最初のままで処理するのではなく、目的が同じ共通点を持つ仲間の式に変形することを示すよい指導例でもある。＞

回帰直線の意味は、2 種類で一方の原因が他方に結果を出させているような関係が見えるときに有効な見える化の一つである。実際に実行してみると、

標準化されたデータ第i番目の組を zw 座標平面上の点と見做した$\mathrm{P}_i(z_i,w_i)\ (i=1{,}2{,}3\cdot\cdot\cdot,n)$と直線$w=rz$と w 軸方向の乖離である誤差（ギャップ）を$\varepsilon_i$とするとき、この誤差の平方の和が最小となる傾き$r$を求める（これを‘最小 2 乗法’という）ことと、誤差の平方の平均が最小となるような回帰直線の傾きrとは同じである。それは、$\varepsilon_i=w_i-rz_i$であり、$\sum_{i=1}^{n}{\varepsilon_i}^2=\sum_{i=1}^{n}(w_i-rz_i)^2$ の最小性は、$\frac{1}{n}\sum_{i=1}^{n}{\varepsilon_i}^2=\frac{1}{n}\sum_{i=1}^{n}(w_i-rz_i)^2$ としても縦軸方向の縮小なので最小性は不変である。これを支える「カバリエリの原理」の指導も大切である。
そこで、$\overline{\varepsilon^2}=\frac{1}{n}\sum_{i=1}^{n}(w_i-rz_i)^2$ とおけるが、これは、
$\overline{\varepsilon^2}=\overline{(w-rz)^2}$ のことであるから、
$\overline{\varepsilon^2}=\overline{w^2-2rzw+r^2z^2}=\overline{w^2}-2r\overline{zw}+r^2\overline{z^2}$ $\cdots*$
ここで、z,wは標準化されているので、
分散はそれぞれ$\overline{w^2}=1,\overline{z^2}=1$であり、

$\overline{zw}=\overline{\left(\frac{x-\bar{x}}{\sigma_x}\right)\left(\frac{y-\bar{y}}{\sigma_y}\right)}=\frac{\overline{(x-\bar{x})(y-\bar{y})}}{\sigma_x\sigma_y}=\frac{\sigma_{xy}}{\sigma_x\sigma_y}$ であるから、

式＊は、
$\overline{\varepsilon^2}=r^2-2\frac{\sigma_{xy}}{\sigma_x\sigma_y}+1=\left(r-\frac{\sigma_{xy}}{\sigma_x\sigma_y}\right)^2+1-\left(\frac{\sigma_{xy}}{\sigma_x\sigma_y}\right)^2$ $\cdots$★
この$\overline{\varepsilon^2}$ の最小値は、$r=\frac{\sigma_{xy}}{\sigma_x\sigma_y}$ のときである。
言い換えると、相関係数は、標準測度の積の平均である。

＜イ＞ $-1\leqq r\leqq 1$について

$\overline{\varepsilon^2}=r^2-2\frac{\sigma_{xy}}{\sigma_x\sigma_y}+1\geqq 0$ であるから、

★の最小値≧０⇔判別式＝$\left(\frac{\sigma_{xy}}{\sigma_x\sigma_y}\right)^2-1\leqq 0$
すなわち　$(r)^2-1\leqq 0$　よって、$-1\leqq r\leqq 1$　であることもわかる。

＜ウ＞　誤差の分散の最小値

誤差 $\varepsilon_i=w_i-rz_i$の平均 $\bar{\varepsilon}=\overline{w-rz}=0$であるから、式★は、誤差$\varepsilon$の分散

$$\sigma_\varepsilon{}^2=\overline{(\varepsilon-\bar{\epsilon})^2}=\overline{\varepsilon^2}=\left(r-\frac{\sigma_{xy}}{\sigma_x\sigma_y}\right)^2+1-\left(\frac{\sigma_{xy}}{\sigma_x\sigma_y}\right)^2$$

であるから、$r=\frac{\sigma_{xy}}{\sigma_x\sigma_y}$のとき、誤差の分散$\overline{\varepsilon^2}$ が最小となることを示している。これがデータの間を走る回帰直線として最もフィットする(傾き)といえる。

＜最小２乗法という表現は、計算方法であり、目標を表現してはいないので教育的ではない。回帰直線とは、データの点の間を走る代表直線で、「その傾きは、誤差(データの点と直線との縦軸方向のギャップ)の分散が最小となるとき」、が教育的であろう。＞
これより、平均$(\bar{x},\bar{y})$を原点(0, 0)に平行移動し標準偏差で割った標準化された座標系（正規直交座標系）において、回帰直線の傾き(回帰変数)＝相関係数である。

$$\frac{y-\bar{y}}{\sigma_y}=r\frac{x-\bar{x}}{\sigma_x}\qquad・・・①$$

＜エ＞　標準化以前の散布図おける回帰直線の傾き(回帰係数)はどうなるのだろうか

$\frac{y-\bar{y}}{\sigma_y}=r\frac{x-\bar{x}}{\sigma_x}$…①は平均値の点$(\bar{x},\bar{y})$を通る直線である。

では、点$(\bar{x},\bar{y})$を通る直線 $y-\bar{y}=a(x-\bar{x})$…② の傾きaは何だろうか。実は簡単で、
①から$y-\bar{y}=r\frac{\sigma_y}{\sigma_x}(x-\bar{x})$　これが②であるから

$$a=r\frac{\sigma_y}{\sigma_x}$$

xを説明変数、yを目的変数と呼ぶことがある。xが原因でyがその結果というときに、xがyに与える影響を直線で見える化したものである。

このように、たとえば、薬の用量xとその効果yの関係のように、２つのデータ間に明確な因果関係があるときを想定したときに用いるものである。

$\frac{y-\bar{y}}{\sigma_y}=r\frac{x-\bar{x}}{\sigma_x}$　・・・①を再考してみよう：

単純に平均を取ると、 0＝0 となり、新たな式はでてこない。そこで、①の両辺に$x-\bar{x}$を掛けて変形した、
$(x-\bar{x})(y-\bar{y})=r\frac{\sigma_y}{\sigma_x}(x-\bar{x})^2$の両辺の平均を考える。
この式の平均は、$\overline{(x-\bar{x})(y-\bar{y})}=r\frac{\sigma_y}{\sigma_x}\overline{(x-\bar{x})^2}$　これは、$\sigma_{xy}=r\frac{\sigma_y}{\sigma_x}\sigma_x{}^2$　ゆえに、$r=\frac{\sigma_{xy}}{\sigma_x\sigma_y}$

つまり、最小２乗法を使わずに‘平均化’で導出可能であることが分かったので指導しやすい。

＜オ＞　標準化していない散布図での回帰直線と最小 2 乗法

＜従来の方法＞ 最小２乗法で、$\sum_{i=1}^{n}\{y_i-(ax_i+b)\}^2$を最小にするような$a,b$を平方完成または偏微分で求め、$a=\frac{\sigma_{xy}}{\sigma_x{}^2}$, $b=\bar{y}-a\bar{x}$ であることを示す。結果、回帰直線は平均点$(\bar{x},\bar{y})$を通過することも示せる。しかし、従来のままでは、Σを使わないと式が長くなり、平方完成も辛くなる。だが、これも、データ第i番目の組を xy 座標平面上の点と見做した $\mathrm{P}_i(x_i,y_i)$ $(i=1,2,3・・・,n)$と、直線$y=ax+b$と y 軸方向の誤差（ギャップ）をε_iとし、
$\sum_{i=1}^{n}\varepsilon_i{}^2=\sum_{i=1}^{n}\{y_i-(ax_i+b)\}^2$ …＃
ここで、$\overline{\varepsilon^2}=\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i{}^2$とおくと、＃の最小値ではなく＃の平均である $\overline{\varepsilon^2}$=$\overline{\{y-(ax+b)\}^2}$ の最小値を調べればよい。理由は平均をとっても最小性は不変である。

$\overline{\varepsilon^2}-(\bar{\varepsilon})^2$＝$\sigma_\varepsilon{}^2$と分散との関係が気になる。
これは、$\overline{\varepsilon^2}$＝$\sigma_\varepsilon{}^2+(\bar{\boldsymbol{\varepsilon}})^2\geqq\sigma_\varepsilon{}^2$ …♭となり、
等号は、$\bar{\boldsymbol{\varepsilon}}=\bar{y}-a\bar{x}-b=0$ のときであるとわかり、
結局、$\overline{\varepsilon^2}$の最小は、$\sigma_\varepsilon{}^2$が最小となるa,bを求めればよい。
$\varepsilon-\bar{\varepsilon}=\{y-(ax+b)\}-\{\bar{y}-a(x-\bar{x})\}$ であるから、
$\sigma_\varepsilon{}^2=\overline{(\varepsilon-\bar{\varepsilon})^2}=\overline{\{(y-\bar{y})-a(x-\bar{x})\}^2}$
$=\overline{(y-\bar{y})^2-2a(x-\bar{x})(y-\bar{y})+a^2(x-\bar{x})^2}$
$=\sigma_y{}^2-2a\sigma_{xy}+a^2\sigma_x{}^2=\sigma_x{}^2\left\{\left(a-\frac{\sigma_{xy}}{\sigma_x{}^2}\right)^2-\frac{\sigma_{xy}{}^2}{\sigma_x{}^4}\right\}+\sigma_y{}^2$
$=\sigma_x{}^2\left(a-\frac{\sigma_{xy}}{\sigma_x{}^2}\right)^2-\frac{\sigma_{xy}{}^2}{\sigma_x{}^2}+\sigma_y{}^2$
これを最小にするaは、$a=\frac{\sigma_{xy}}{\sigma_x{}^2}$ であることがわかった。これが、回帰直線の傾き(回帰係数) aであり、♭の等号成立が $\bar{\varepsilon}=\bar{y}-a\bar{x}-b=0$ のときより、切片は $b=\bar{y}-a\bar{x}$ となり、回帰直線は平均点$(\bar{x},\bar{y})$を通過するという重要事項も同時に示せる。