

学校数学としての統計の指導について

—— 散布度の扱いを中心にして ——

藤 本 義 明

(数学科教育研究室)

(平成10年9月30日受理)

On the teaching of statistics as the school mathematics :focus on the teaching of dispersion

Yoshiaki FUJIMOTO

I. はじめに

21世紀にむけての学習指導要領の改正の動きが顕著になってきた。教育課程審議会の「審議のまとめ」が本年6月に出され、学習指導要領の改正の方向が示された。これによって、新しい学習指導要領の骨格が見えてきたところである。

新しい学習指導要領の方向の中で、中学校数学における統計の扱いは大きく様変わりしそうである。これまで、「資料の整理」として中学校数学で常に扱われてきた統計の指導は、中学校数学から消える可能性が出てきた。教育課程審議会の「審議のまとめ」によると、中学校での「資料の整理」は高等学校で一括して扱うということである。

学校数学の中で、「資料の整理」または統計というのは、社会的な有用性がとくに大きな教材として位置付けられてきた。したがって、小学校や中学校の義務教育段階でつねに社会とのかかわりが深いことを大きな理由として指導されてきたはずである。このような性格を持つ「資料の整理」または統計を、教科の専門性がより強い高等学校数学でまとめて指導するという考え方は理解に苦しむところである。社会生活に関連が深い教材を義務教育段階では軽視してより専門的段階で指導するというのは筋の通らない話である。考えられることは、今までの「資料の整理」または統計の位置付けが誤っていたということなのかもしれない。確かに、これまでの「資料の整理」の指導には、いくつかの問題があったと思われる。そこで、本稿では、中学校、高等学校における「資料の整理」または統計の指導を検討してみる。

II. 問題の所在

1. 標準偏差

「資料の整理」でもっとも問題なのは散布度の扱いであろう。現在の数学教育では、散布度として、中学校では「範囲（レンジ）」、高等学校では「標準偏差」を指導することになっている。問題は、範囲は散布度として簡単過ぎてあまり有用性がないこと、逆に、「標準偏差」はその意味が生徒にはまったく不明であるということである。とくに、「標準偏差」は各資料の偏差（平均値との差）の2乗によって分布の散らばりを示そうとしているわけだが、偏差の2乗が分布の散らばり度とどのように関係するのか、生徒には理解できない。考えてみると、小学校、中学校、高等学校の数学の内容において、どんなに難しい概念でも、教師がその意味を説明できないような概念は存在しない。どんなに難しい概念でも、教科書には生徒がなんとか理解できるような説明がなされているものである。しかし、「標準偏差」はなかなか説明が困難で、教科書にも説明なしに登場する概念である。したがって、生徒は標準偏差の意味はわからず、その使い方だけを丸暗記する結果になっている。

2. 相関係数

相関係数も標準偏差を用いているので、その意味がわかりにくい。その上、相関係数の意味を説明している教科書はほとんどない。例えば、次のような記述がせいぜいである。

『 2種類の変量 x , y をもつ統計資料を

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

とし、 x , y の平均を \bar{x} , \bar{y} , 標準偏差を s_x , s_y で表す。

また、 x , y のそれぞれの偏差の積の平均

を x , y の共分散といい、 s_{xy} と表す。

s_{xy} を $s_x s_y$ で割ったものを x , y の相関係数といい、 r で表す。

つまり、 r は次のようになる。

$$r = \frac{\frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})(y_k - \bar{y})}{\sqrt{\frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2} \sqrt{\frac{1}{n} \sum_{k=1}^n (y_k - \bar{y})^2}} \quad \text{J}^{\text{①}}$$

共分散の意味は相関表と関連させながら、一応の説明はなされている。しかし、共分散を標準偏差の積で割ることの意味や必要性を説明している教科書はほとんど無い。したがって、生徒にとって、相関係数は理解不能である。

3. 正規分布

正規分布は推定・検定に用いられる基本的な連続分布であるが、教科書におけるその出所はきわめてあいまいである。例えば、ある教科書^②では正規分布について次のような説明で終わっている。

* $y = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$ で表される曲線を正規曲線という。

*身長分布、測定誤差などの分布は正規分布に近い。

*二項分布 $B(n, p)$ は、 n が大きくなるにつれて正規分布に近づく。

このような説明からは、 $y = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$ という分布の式が与えられる理由は生徒にはわからない。とくに、この式の定義の中にあらかじめ平均 m や標準偏差 σ が現れていることなどは生徒を混乱させる。通常は分布を調べた後で、その平均や標準偏差が知れるはずだから、はじめから平均や標準偏差がわかっているかのように思われるこの式は理解しにくい。

4. 推 定

母平均の推定について、教科書では次のような記述が見られる。

母平均 m 、母標準偏差 σ の母集団から、大きさ n の標本を無作為抽出し、その標本平均を \bar{X} とする。 n が大きいとき、標本平均 \bar{X} の分布は正規分布 $N\left(m, \frac{\sigma^2}{n}\right)$ に近いから、 \bar{X} を標準化した確率変数 $Z = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$ の分布は標準正規分布 $N(0, 1)$ に近い。

このあと、正規分布表から

$$\text{確率 } P\left(-1.96 < \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} < 1.96\right) = 0.95$$

$$\text{確率 } P\left(\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}} < m < \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

同様に

$$P\left(\bar{X} - 2.58 \times \frac{\sigma}{\sqrt{n}} < m < \bar{X} + 2.58 \times \frac{\sigma}{\sqrt{n}}\right) = 0.99$$

を導いて、次のようにまとめている。

『以上のことから、次のように母平均を推定することができる。』

母平均の推定

母標準偏差 σ の母集団から大きさ n の標本を無作為に抽出するとき、 n が大きいならば、母平均 m に対する信頼区間は

(1) 信頼度95%では

$$\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}} < m < \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}$$

(2) 信頼度99%では

$$\bar{X} - 2.58 \times \frac{\sigma}{\sqrt{n}} < m < \bar{X} + 2.58 \times \frac{\sigma}{\sqrt{n}}$$

標本の標準偏差を、標本標準偏差といい、 s で表す。実際の標本調査では、母集団の標準偏差 σ がわからない場合が多い。しかし、 n が大きいときは、 σ の代わりに標本標準偏差 s で代用してよいことが知られている。』^③

このような説明に基づいて、実際の推定では、標本標準偏差 s を使って、母平均の推定をさせている。これに対して、ある生徒が次のような質問をした。「母標準偏差の変わりに標本標準偏差を用いてよいならば、母平均の変わりに標本平均を用いてもよいではないか。」

この生徒の反応は実に素直なものであり、教科書の記述のように、標準偏差だけ母標準偏差の変わりに母標本標準偏差を代用するというのは、筋が通らない。このような記述はこの教科書に限ったことではなく、すべての教科書がこのような記述をしている。

以上のように、「資料の整理」や統計の指導では、標準偏差が扱われる場面以降、他の分野では決して見られないような、生徒の理解とはかけ離れた指導場面が現出しているのが実情である。

Ⅲ. 散布度の指導

「資料の整理」または統計の指導の一番のネックは、散布度としての「標準偏差」である。なぜなら、標準偏差は、相関係数、推定・検定のための基本的概念であるから、標準偏差が理解できないとしたら、統計の理解はあり得ない。本稿では、散布度として、意味のわかりにくい標準偏差ではなく、意味付けが容易な平均偏差を導入することを提案する。

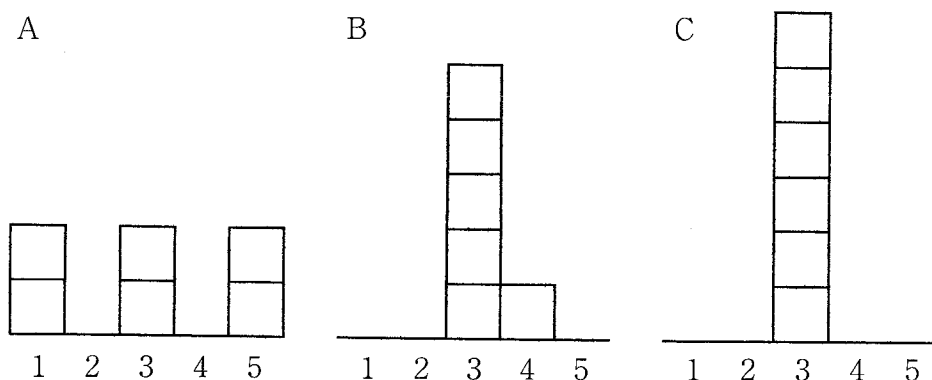
まず、その前提として、生徒の散らばり具合についての捉え方の実態を調べてみる。

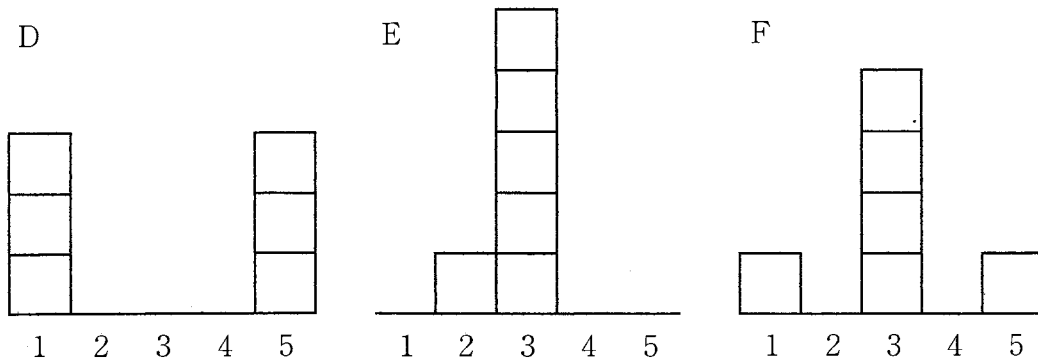
1. レディネスの調査

散布度のレディネスを調査するために、散布度を未習である中学校1年生38名を対象に、次のようなアンケートを実施した。

(問題) 班員が6人ずつの班A～Fがあり、5点満点の数学のテストをしたところ、各班の点が下のグラフのようになった。テストの点は各班によって違いがある。みんなの点がバラバラにばらけている班もあるし、みんなの点がまとまっている班もある。各班でのテストの点のまとまりぐあいについて、まとまりのある班から順に答えなさい。

注意：グラフの中で、1ます が生徒1人を表している





まとまりのある順番に班の名前を下のカッコの中にかきなさい。ただし、同じまとまりぐあいであると思ったら、1つのカッコの中に2つ以上の班の名前を書いてもよい。

<答え>
 まとまりのある順番に
 () → () → () → () → () → ()

<結果>

Cを最下位にしている者が6人いたが、これらの者は、問題の意味を間違えて、まとまりの少ない順番に答えと考えられる。したがって、この6人の答えは逆にする。

また、BとEは同じであると答えた者は5人であった。この5人の回答はBEの順で統一する。すると、38人の回答は次のように整理できる。

CABEDF・・・1人	CBEDAF・・・5人	CDABEF・・・2人
CADBEF・・・1人	CBEDFA・・・4人	CDAFBE・・・2人
CADEBF・・・1人	CBEFAD・・・2人	
CADFBE・・・3人	CBEFDA・・・7人	
CEBDAF・・・1人	CFDABE・・・1人	
CEBDFA・・・2人	CFDEBA・・・2人	
CEBFDA・・・1人		
ACDFBE・・・1人		
ADCBFE・・・1人		
ADCFBE・・・1人		

<分析>

①BEが並んでいるもの 32人

BとEが同じ散らばり具合であったとした者5人と、BEを並べて書いた者33人を合わせるとほぼ100%の占有率となる。つまり、左右対称をおなじ散らばり具合と見ることは、ほぼ完成

している。

②BまたはEが2番目の者 22人 58%

Cを1番にした者92%と合わせて考えると、「散らばりが全くない」とか「ほとんどない」という状態の理解はほぼ完成しているといえるだろう。

③最後 A・・・16人, BまたはE・・・9人,
D・・・2人, F・・・・・・・・・・11人

最後、つまり最も散らばっているとみなされた状態をDとした者は少ない。これは、山の数が少ないのでかたまっていると判断し、最後にはしなかったということであろう。

また、AやFがかなり多い。これは、山が多いので散らばっていると判断したということであろう。

したがって、生徒の散らばり具合の判断には、山の数が大きなウェートを占めている。ただし、山の数が等しいAとFが必ずしも隣り合っていないから、単純に山の数だけで判断しているのでもない。生徒は範囲と山の数を複合させた見方をしているようである。

④				累計
正解	CBFAD	2	5%	
ほぼ正解	CBFDA	8	21%	
正解に近い	CB DFA	6	16%	42%

全くの正解者は極めて少ない。正解に近い者の累計は39%であるから、レディネスがあるとは言えないだろう。散布度の理解は、授業のやり方次第であるということだろう。

2. 歴史的背景

平均偏差と標準偏差を比べた場合、有用性が高いのは標準偏差であるが、理解のしやすさから見ると、明らかに平均偏差が理解しやすい。それでは、どうして、わかりやすい平均偏差ではなくて、きわめてわかりにくく標準偏差が教えられているのだろうか。それは単に有用性という理由だけのためであろうか。また、標準偏差ではなくて平均偏差が教えられたことは今までにないのだろうか。このような疑問を解決するために、統計学の歴史や統計教育の歴史を振り返ってみる。

(1) 統計学史的背景

小杉氏は、統計の考えの起源は、古代国家の時代にまでさかのぼるといふ。^④氏は古代国家における統計の起こりを説明したのち、おおむね、次のような発達歴史区分を行っている。

- | | | |
|-------------|------------------|------------------|
| ① ドイツ大学派統計学 | ② 政治算術 | ③ 古典確率論 |
| ④ 初期人口論 | ⑤ ケトレー統計学 | ⑥ 記述統計学としての数理統計学 |
| ⑦ 社会統計学 | ⑧ 推測統計学としての数理統計学 | ⑨ 数理統計学 |

各時代の概略は次のようである。

① ドイツ大学派統計学 (16世紀～19世紀)

今日の統計学という名称のもとになった「国家学」Statistik という学問が栄えた。方法的には国家の情勢を文章で記述するのが主であり、現在の統計学に直接つながるものではなかった。

② 政治算術 (17世紀～18世紀)

イギリスを中心に、商人によってドイツ大学派統計学とは独立に発展した。社会や経済を知

るために、現在の統計学の手法につながる「大量観察の方法」が用いられて数量的観察と分析が行なわれたのが大きな特徴である。

③ 古典確率論（16世紀中葉～）

統計学と関連が深く、統計学を理論的に支える確率論が発達した。ガウスの最小自乗法（1795）、ポアソンの大数の法則（1837）など、めざましい発展を遂げた。

④ 初期人口論（17世紀～19世紀）

マルサスの人口論など、人口増加の予測的研究は、統計学の内容的動機を与えた。

⑤ ケトラー統計学（19世紀初頭～20世紀初頭）

現代統計学の直接のルーツになる統計学が発達した時期である。この時期の代表であるケトラーは、確率論に基づいて、統計学に大数法則という基礎を与え、様々な社会現象を研究した。例えば、身長がガウスの誤差分布（ガウス曲線、正規曲線）になることなどを発見した。

⑥ 記述統計学としての数理統計学（19世紀後半～）

統計方法の研究が重んじられ、記述統計の方法がいろいろと開発された。例えば、Galton は、回帰直線、相関関係、中位数、四分位数、四分位偏差などを創案した。そして、Pearson は、1893年に、平均偏差と標準偏差、1894年にモード、1897年に重相関、1900年に X^2 テストなどを創案した。

⑦ 社会統計学（19世紀後半～20世紀初頭）

統計学を実質科学と見て、国勢調査のような大量観察法に限定した統計学を研究した。

⑧ 推測統計学としての数理統計学（19世紀末～）

統計的方法だけを抽出して、その学問的体系付けをしようとする傾向が生じてきた。Fisher を代表に、標本理論を中心とした現代数理統計学の方向が定まった。

⑨ 数理統計学（20世紀中葉～）

現代の数理統計学が確立された。

<分析>

統計学の歴史でみると、標準偏差と平均偏差がともに、Pearson によって考え出されたことがわかるが、これは意外なことである。標準偏差と平均偏差とでは概念の難しさにおいて差があるように思われるから、年代的に平均偏差が先に現れ、その後で標準偏差が出て来るのが自然なところだろう。しかし、実際には、両者は同一の機会で見られたのである。また、四分偏差やモードなどの素朴なアイデアや、重相関や X^2 テストといった高度なアイデアが入り混じって考案されたのであった。しかも、それは19世紀末のことであって、すでに、確率論はかなりの発展をした後であり、解析学などもすでに19世紀には目覚ましい発展をした後のことである。

すなわち、今日中学校で教えられているような統計学の基礎概念は、数学が高度に発展した後には生まれたため、文化的・社会的な成熟をする機会がないまま、高度な数学の理論に飲み込まれて行ったと考えられる。したがって、これらの概念には理解しやすいかどうかという基準はあてはまりにくく、数学理論的に有用であるかどうかという基準に支配されているのであろう。

(2) 教育史的背景

統計学の歴史から見て、平均偏差が標準偏差にとって変わることの難しさがあることは明らかになったが、わが国の数学教育の歴史の中では、これらはどのように扱われてきたのだろうか。過去に平均偏差が指導された実績というのではないのだろうか。

ここで、戦後の数学教育を振り返って、学習指導要領の範囲で、散布度の扱いの変遷を見てみる。^⑤ ただし、表中の改定年度は中学校学習指導要領の改訂年度で示し、次の略号を用いる。

(略号) S S・・・「資料の整理での標準偏差」を扱う

K S・・・「確率分布における標準偏差」を扱う

S K S・・・上記の2つとも扱う

改定年度	中学校	高等学校	
昭和35	なし	(数学ⅡA) S S	(数学Ⅲ)主として, K S
昭和45	主として S S	(数学ⅡA) K S	(数学Ⅲ) K S
昭和52	範囲	(数学Ⅱ) S K S	(確率統計) S K S
平成元	範囲	(数学B) K S	(数学C) S K S

昭和35年と昭和45年に見られる「主として資料の整理での標準偏差」, 「主として確率分布における標準偏差」というのは、散布度として平均偏差を扱うことにも含みを持たせていたとも思われるが、それも昭和40年代までで、昭和52年の改訂以降は、中学校では「範囲」、高等学校では「標準偏差」が定着しており、平均偏差はまったく考えられていない。

生徒にとっては確率分布自体が分かりにくいので、確率分布における標準偏差は余計にわかりにくい。標準偏差を導入するには、まず資料の整理において導入すべきである。現在の高等学校数学Bではこのような配慮はされでならず、いきなり、確率分布で標準偏差を扱っている。生徒の標準偏差の理解という問題にいかに関心であるかを表している。

Ⅳ. 指導の可能性

散布度として標準偏差の変わりに平均偏差を導入したいわけだが、たとえ生徒の理解にとっては平均偏差が有利であっても、統計の系統的学習にまったく使えないのでは意味がない。ここでは、平均偏差が統計の系統的学習にどこまで使用できるかを実証してみたい。そのために、現行の高等学校の数学Cの教科書を例として、そこで使われている標準偏差を平均偏差に置き換えてみる。ここで用いる教科書は、第一学習社『高等学校 数学C』であるが、他の教科書でも標準偏差の扱いに大きな違いはないので、この教科書を例とする。

p.104～p.146 第4章 統計処理

§1 統計資料の整理

① 代表値と散布度

p. 107

例4 (与えられた資料の) 小テストについて、A組の標準偏差を求める。

得点	0	1	2	3	4	5	計
A組	0	0	4	13	18	5	40
B組	0	0	9	8	13	10	40

平均は3.6で (標準偏差) = $\sqrt{0.69} \doteq 0.831$

<平均偏差> = 0.71

p. 108

問5 (上の資料の) 小テストについて, B組の得点の標準偏差を求めよ。

平均は3.6で (標準偏差) = $\sqrt{1.19} = 1.091$
 <平均偏差> = 0.96

p. 109

例5 (簡便法により) A組の標準偏差を求める。

(標準偏差) = $\sqrt{0.69} \doteq 0.831$
 <平均偏差> = 不能

問6 (簡便法により, 与えられた) 身長の数値分布表から, 標準偏差を求めよ。

身長

階級値	160	164	168	172	176	180	計
度数	1	3	5	2	6	3	20

(標準偏差) = $\sqrt{35.04} \doteq 5.919$
 <平均偏差> = 不能

練習3 (簡便法により, 与えられた) 体重の数値分布表から, 標準偏差を求めよ。

階級値	50	54	58	62	66	70	74	78	82	計
度数	2	4	5	1	1	2	1	3	1	20

(標準偏差) = $\sqrt{102.56} \doteq 10.127$
 <平均偏差> = 不能

② 相関

p. 112

例7 資料について, 相関係数を求める。

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

(xの平均) = 4, (yの平均) = 5
 (標準偏差を利用すると) = 0.76
 <xの平均偏差> = 2, <yの平均偏差> $\doteq 2.67$ より

xの偏差値	-1.5	0.5	-1	1.5	1	-0.5
yの偏差値	1.5	0.37	-0.37	-1.12	-1.5	1.12

<平均偏差を利用すると> = -0.91

問9 (与えられた) 資料について, 例7と同じ方法で相関係数を求めよ。

x	7	1	4	8	6	4
y	5	2	4	9	8	8

(標準偏差を利用すると) = 0.72
 <平均偏差を利用すると> = 0.89

p.113

問10 (与えられた) 資料について, 上の式 (筆者注: 簡便法の式) を用いて相関係数を求めよ。

x	1	3	5	8	7	9
y	5	1	4	5	9	3

(標準偏差を利用すると) = 0.23

<平均偏差を利用すると> = 不能

§ 2 統計的な推測

① 母集団と標本

② 母集団分布と標本平均の分布

p.123

例 4 4枚の硬貨を投げたとき, 表が出る枚数をXとする。

平均2より, (分散) = 1

$$\langle \text{平均偏差} \rangle = \sqrt{2 \times \frac{1}{16} + 1 \times \frac{4}{16} + 0 \times \frac{6}{16} + 1 \times \frac{4}{16} + 2 \times \frac{1}{16}} = \frac{\sqrt{3}}{2} \doteq 0.866$$

問 5 白球 8 個, 黒球 2 個がはいっている箱から, 同時に 5 個の球をとり出すとき, その中に含まれている白球の個数 X の平均と標準偏差を求めよ。

X	3	4	5	計
確率	$\frac{2}{9}$	$\frac{5}{9}$	$\frac{2}{9}$	1

平均 4 より, (標準偏差) = $\frac{2}{3}$

$$\langle \text{平均偏差} \rangle = 1 \times \frac{2}{9} + 0 \times \frac{5}{9} + 1 \times \frac{2}{9} = \frac{4}{9} = 0.44$$

p.124

問 6 (与えられた資料において) 母標準偏差を求めよ。

確率変数 X	1	2	3	計
確率	0.4	0.5	0.1	1

(母平均) = 1.7 より (母標準偏差) = 0.64

$$\langle \text{母平均偏差} \rangle = 0.7 \times 0.4 + 0.3 \times 0.5 + 1.3 \times 0.1 = 0.56$$

問 7 全校生徒 960 名の高校で, ある意見について賛否を調査したところ, 賛成が 640 名で, 反対が 320 名であった。全校生徒を母集団として, 生徒 1 名を無作為抽出し, その生徒の意見が賛成なら 1, 反対なら 0 の値をとる変数 X を考える。このとき, X の確率分布を求めよ。また, 母平均と母標準偏差を求めよ。

X	0	1
確率	$\frac{2}{3}$	$\frac{1}{3}$

(母平均) = $\frac{1}{3}$ より

$$\begin{aligned} (\text{母標準偏差}) &= \sqrt{\frac{2}{9}} \doteq 0.47 \\ \langle \text{母平均偏差} \rangle &= \frac{1}{3} \times \frac{2}{3} + \frac{2}{3} \times \frac{1}{3} = \frac{4}{9} \doteq 0.44 \end{aligned}$$

p.126

例7 球に書かれた数字をXとすると、母集団分布、すなわち、Xの確率分布は右のようになる。

X	1	2	3	4	5	計
確率	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	1

したがって、母平均 m 、母標準偏差 σ は

$$\begin{aligned} m &= 3, \quad \sigma = \sqrt{2} \\ \langle \text{母平均偏差 } h \rangle &= \frac{6}{5} \end{aligned}$$

また、(前ページの) \bar{X} の確率分布表

X	1	1.5	2	2.5	3	3.5	4	4.5	5	計
確率	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{3}{25}$	$\frac{4}{25}$	$\frac{5}{25}$	$\frac{4}{25}$	$\frac{3}{25}$	$\frac{2}{25}$	$\frac{1}{25}$	1

から、標本平均 \bar{X} の平均 $E(\bar{X})$ と標準偏差 $\sigma(\bar{X})$ を求めると

$$E(\bar{X}) = 3, \quad \sigma(\bar{X}) = 1$$

$$\langle \text{平均偏差} \rangle \bar{h}(X) = \frac{4}{5}$$

したがって、次の関係が成り立っている。

$$E(\bar{X}) = m \quad \sigma(\bar{X}) = \frac{\sigma}{\sqrt{2}}$$

*ただし、 h と $h(\bar{X})$ の関係は見出せない。平均偏差を扱う限界が近い。(筆者注)

一般に、次のことが知られている。

母平均 m 、母標準偏差 σ の母集団から、大きさ n の標本を復元抽出するとき、標本平均 \bar{X} の平均 $E(\bar{X})$ 、標準偏差 $\sigma(\bar{X})$ は

$$E(\bar{X}) = m, \quad \sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

問9 母平均55、母標準偏差12の母集団から、大きさ100の標本を復元抽出するとき、標本平均 \bar{X} の平均と標準偏差を求めよ。

$$(\text{標準偏差}) = \frac{12}{\sqrt{100}} = 1.2$$

$\langle \text{平均偏差} \rangle = \text{標準偏差を前提とした問題であるから不能}$

③ 正規分布とその利用

④ 統計的推測の考え

以下、正規分布の理論や正規分布を利用した統計的推測についての教材である。正規分布では標準偏差のかわりに平均偏差を用いるのは困難である。

<分析>

標準偏差のかわりに平均偏差を使うことにより、正規分布が利用される場所までは指導で

きることがわかる。正規分布が利用されるのは、推定や検定だけであるから、現在の統計指導のかなりの部分は平均偏差で間に合うといえる。散布度の意味付けは平均偏差で行ない、相関係数で利用したあと、推定・検定の指導において平均偏差の代わりに標準偏差を使うという指導の流れを推奨したい。

なお、標準偏差や相関係数の簡便法の指導では、平均偏差の使用は不能であった。しかし、簡便法は計算機がなく手計算で計算する時代には重要であったかもしれないが、現在の計算機が発達した時代にはあまり必要なものでもない。したがって、簡便法で平均偏差が使えないことは無視して良いと考える。

V. 提 案

1. 平均偏差

本稿での提案の第一は散布度として、標準偏差の代わりに平均偏差から導入するということがであったが、ここで、平均偏差を導入する教授方法についても提案しておく。生徒は散らばりのレディネスとして、必ずしも平均偏差を意識できていないという状況に対して、平均偏差の導入は決して楽ではない。ある程度恣意的な場面を提示して、平均偏差によって散らばりを捉える意識を形成しなければならない。そのためには、恣意的なモデルを提示することが有効であると考えられる。ここでは、そのような2つのモデルを提案する。

(1) 逃亡のモデル

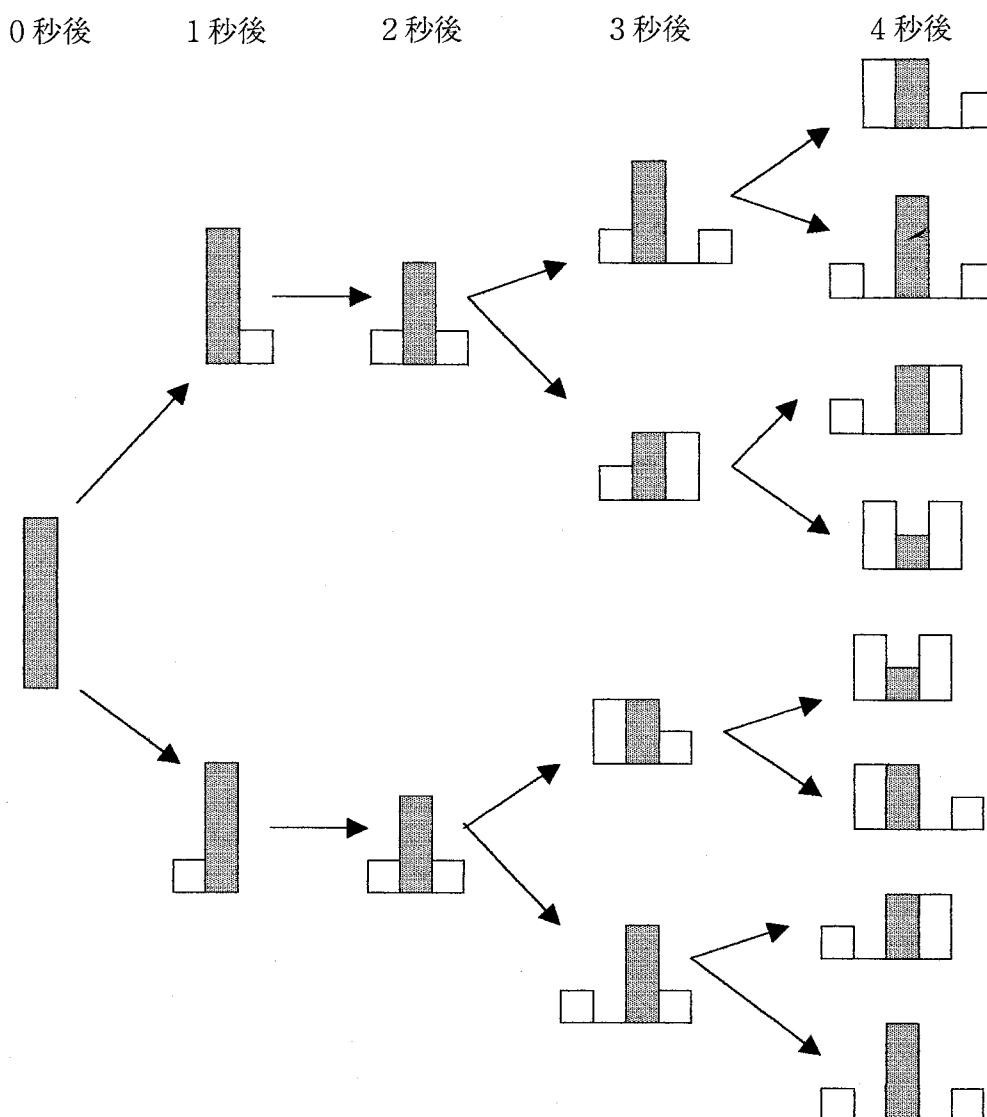
分布がいくつの山に分かれるかという認識ではなく、中心からの偏差を観点として、散らばりを大小を意識させるためのモデル

<モデル>

ある地点に固まっている5人の人が、次のように左右に散らばって行く。

- ① 1秒毎に一人だけが1マスだけ動く。
- ② 右に動く人と左に動く人は決まっている。右に動く人は右にだけ、左に動く人は左にだけしか動けない。
- ③ 右と左に交互に動いていく。

このとき、4秒後までの動きは、次頁の図のようである。



問1：次のア～エは何秒後のようすであるか。

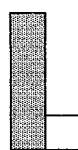
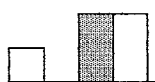
問2：また、これらの散らばり具合の順番はどうか。

ア.

イ.

ウ.

エ.



(問1の答え) ア. 4秒後 イ. 6秒後 ウ. 1秒後 エ. 8秒後

毎秒、外へ動いて行くことで、時間がたつほど散らばりは大きくなると考えられると説明する。

(問2の答え) 散らばり度は、ウ、ア、イ、エの順に大きくなる。

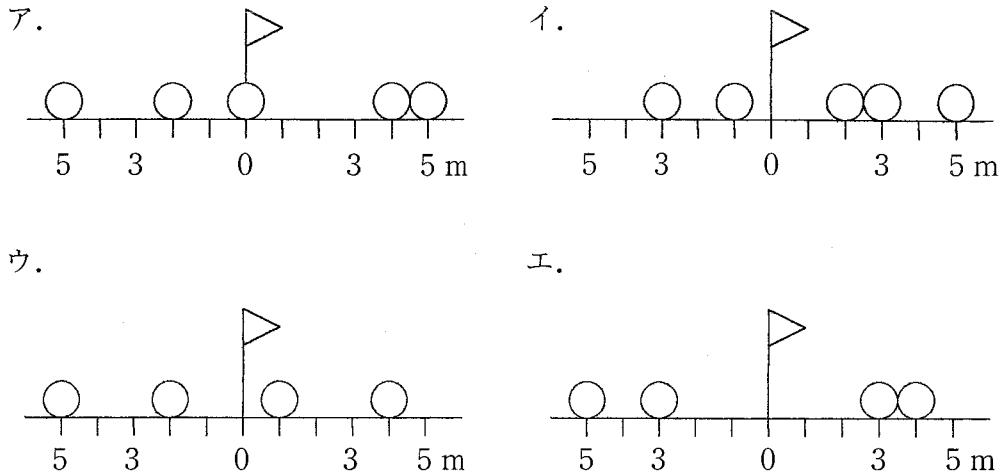
(2) ゴルフのモデル

平均偏差の求め方を意識させるためのモデル

<モデル>

4チームがゴルフのパターを競い合う。あるところからホール（カップ）をねらい、その結果が下図のようになった。ただし、旗の位置がホールである。

問 どのチームのスコアーがもっともよいか。



(答え) ホールからの距離の平均は

ア. 3.2m イ. 2.8m ウ. 3.0m エ. 3.75m

よって、スコアーが良いチームは、イ、ウ、ア、エの順である。

平均の距離が、スコアーの良し悪しを決める。スコアーが良いほどホールを中心に散らばりが少ない。ボールの位置が左右対称に近い場合は、分布の中心をホールとみなしやすいが、一般には分布の中心とホールが一致するとは限らないのが難点である。

2. 偏差値と相関係数について

現在の統計の指導では、散布度を求めてもそれを利用する場面は乏しい。散布度を利用して分布を分析場面として、次のような基準化した値を指導したい。この数値は偏差値の基本概念であるから、ここではこれを「偏差値」と呼ぶことにする。ただし、この場合も散布度としては平均偏差を用いる。

$$x_i \text{ の偏差値 } t_i = \frac{x_i - m}{h} \quad (m : \text{平均値}, h : \text{平均偏差})$$

偏差値の指導においては、例えばテストの得点の場合であれば、平均偏差を『平均的な人の偏差』として意味付け、偏差値は『ある人の得点の偏差が平均的な人の偏差の何倍であるかを示す数値』と意味付ければ、生徒の理解は容易になるであろう。

3. 相関係数について

「Ⅱ. 問題の所在」で見たように、現在の相関係数の指導は計算方法の押し付けだけに終わ

っている。相関係数の意味を理解させるためには、上記の偏差値を使えば良い。127ページの平均偏差を用いた相関係数はこれによっている。

n 個の資料について、 i 番目の資料の変量 X の偏差値を t_{Xi} 、変量 Y の偏差値を t_{Yi} とすると、

$$\text{(相関係数)} = \frac{\sum_{i=1}^n t_{Xi} \times t_{Yi}}{n}$$

ただし、この数式の意味付けは、「Ⅱ. 問題の所在」で示した現在指導されている共分散と同様の意味付けで済み、しかも、その平均ということだから理解も容易である。

Ⅵ. おわりに

本稿では、現在の統計の指導について、範囲、標準偏差、相関係数、推定の問題点を指摘し、それらを解決するかなめとして、散布度に平均偏差を用いることを推奨してきた。散布度の指導に関して、統計学の歴史からみた困難さ、教育史的に見た散布度の指導の不適切さなども指摘した。そして、平均偏差を用いても統計の系統的学習上大きな支障が無いこともわかり、平均偏差の指導を容易にするためのモデルも提案した。

しかし、最後に強調しておきたいことは、単に統計の問題というよりも学校数学の問題である。本稿で指摘したように、学校での統計学は理論数学中心に組み立てられ、生徒の理解からかけ離れている。学校数学は理論数学をうすめただけというのは適切ではない。生徒の理解や教育的見地から、理論数学とは独立に考えられなければならない。もう一つ統計学について言えば、統計学は実質科学として実用性を主に生まれ、次第に形式的方法論にその重心を移し、数学の一分野として育ってきた。したがって、大学の統計学であれば数学理論として扱われるのも当然であるが、文化の誕生と切っても切れない関係にある教育の場では、実質科学の部分が大切である。つまり、学校数学の統計学では実用性という面を意識しなければならない。例えば、正規分布が何物かわからなくても正規分布は有用だから指導するというようなことである。この点有用性以上に理論の整合性を尊重する数学とは異なっている。つまり、授業では数学での証明の重要性を指摘した後に、統計の指導では証明を無視して結果の有用性にすぎるといった朝令暮改を余儀なくされている。この点、教授者は困惑するところである。数学とは性格の異なる統計の指導は、学校数学とは切り離し、例えば応用性が重要であるところの情報学などで指導する方が良いと考える。その意味で、本稿の標準偏差の前に平均偏差を指導するという主張は、統計学を数学で指導しなければならない場合に限った次善の策である。

<引用・参考文献>

- ① 教科書「高等学校数学C」啓林館 平成7年 p.157
- ② 教科書「高等学校数学C」第一学習社 平成7年 pp.128-132
- ③ 同上 pp.134-136
- ④ 小杉 肇 『統計学史』恒星社厚生閣, 1984年
- ⑤ 文部省 中学校数学・高等学校数学の各「指導書」