

1 学習内容の説明 ⇒ 2 問題演習 ⇒ 3 振り返り（確認テスト・相互採点・リフレクションの記入）

【内容目標】 散らばりの度合いを表す値を求められるようになる

□分散と標準偏差

●データ全体の特徴を表そう

データの中のいくつかの代表的な値を用いて散らばりの度合いを表す値 ⇒ 四分位数

データの中の すべての値 を用いて散らばりの度合いを表す値 ⇒ ?

◎ 次のことを用いて、平均値の周りに各値がどのように分布しているかを考えてみよう

**偏差** データの各値と平均値  $\bar{x}$  との差のこと。  $x - \bar{x}$  で表す。  
偏差の総和は 0 であるので、偏差の平均ももちろん 0。

**分散** 偏差  $x - \bar{x}$  の 2 乗の平均値のこと。  
変数  $x$  のデータの値が  $x_1, x_2, \dots, x_n$  で、その平均値が  $\bar{x}$  のとき

$$\text{式で表すと} \quad s^2 = \frac{1}{n} \{ (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2 \}$$

**標準偏差(standard deviation)**

分散の正の平方根のこと。  $s$  で表す。要するに  $s = \sqrt{\text{分散}}$ 。  
標準偏差が小さくなるほどデータは平均値の周りに集中しており、散らばりの度合いが小さくなる。逆に標準偏差が大きくなれば散らばりの度合いが大きいといえる（分散も同様である）。

<注意> 標準偏差の単位は、変数  $x$  の単位と同じである。

データの値が平均値の周りに集中しているほど、それぞれの偏差の絶対値は小さくなり、分散、標準偏差も小さくなる傾向にある。

**例 9)** データの分散、標準偏差

10 人の漢字テストの得点  $x$  が、下の表で与えられている。  $x$  の単位は点である。

	$x$	偏差は平均との差を計算 $x - \bar{x}$	偏差の 2 乗 $(x - \bar{x})^2$
①	9	2	4
②	3	-4	16
③	4	-3	9
④	10	3	9
⑤	10	3	9
⑥	5	-2	4
⑦	7	0	0
⑧	9	2	4
⑨	10	3	9
⑩	3	-4	16
総和	70	0	80
平均	7		8

よって、分散  $s^2$  は  $s^2 = \frac{1}{10} \times 80 = 8$

標準偏差  $s$  は  $s = \sqrt{8} \approx 2.8$  (点)

分散にルートを付けたら  
「標準偏差」

数式で書くと（問題集などでは下のようによく書かれることが多い）

$$s_x^2 = \frac{1}{10} \{ (9-7)^2 + (3-7)^2 + (4-7)^2 + (10-7)^2 + (10-7)^2 + (5-7)^2 + (7-7)^2 + (9-7)^2 + (10-7)^2 + (3-7)^2 \} = \frac{1}{10} \times 80 = 8$$

平均を求めるときに仮平均を使ってもよい

「偏差の 2 乗の平均」が「分散」

□分散と平均値の関係式

分散は次のような求め方もできる。

( $x$  のデータの分散) = ( $x^2$  のデータの平均値) - ( $x$  のデータの平均値)<sup>2</sup>  
 つまり、データの値の『2乗の平均値』から、『平均値の2乗』を引いた値が分散に等しい。  
 ただし、データの値が小さくなければ大変なので注意。

<注意>

変数  $x^2$  のデータの平均値を  $\overline{x^2}$  と書くと、上の式は

$s^2 = \overline{x^2} - (\overline{x})^2$  と書ける。

$$\begin{aligned} s^2 &= \frac{1}{n} \{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2\} \\ &= \frac{1}{n} \{(x_1^2 + x_2^2 + \dots + x_n^2) - 2\bar{x}(x_1 + x_2 + \dots + x_n) + n(\bar{x})^2\} \\ &= \frac{1}{n} (x_1^2 + x_2^2 + \dots + x_n^2) - 2\bar{x} \cdot \frac{1}{n} (x_1 + x_2 + \dots + x_n) + (\bar{x})^2 \\ &= \frac{1}{n} (x_1^2 + x_2^2 + \dots + x_n^2) - (\bar{x})^2 \end{aligned}$$

例10) 関係式を用いて、例9のデータの分散を求める。

$x$	9	3	4	10	10	5	7	9	10	3	計 70
$x^2$	81	9	16	100	100	25	49	81	100	9	計 570

$\bar{x} = \frac{1}{10} \times 70 = 7, \quad \overline{x^2} = \frac{1}{10} \times 570 = 57$  である。

よって、分散  $s^2$  は  $s^2 = \overline{x^2} - (\bar{x})^2 = 57 - 7^2 = 8$  (終)

4プロ325) 15個の値からなるデータがあり、そのうちの10個の値の平均値は9、分散は3、残り5個の値の平均値は6、分散は9である。

- (1) このデータの平均値を求めよ。
- (2) このデータの分散を求めよ。

	10個のデータ	5個のデータ
平均値	9	6
総和	$9 \times 10$	$6 \times 5$
15個のデータの総和	$9 \times 10 + 6 \times 5 = 120$	
15個のデータの平均値	$\frac{120}{15} = 8$	
2乗の平均値	$a$	$b$
2乗の総和	$a \times 10$	$b \times 5$
15個のデータでの総和	$a \times 10 + b \times 5$	

(1)  $\frac{9 \times 10 + 6 \times 5}{15} = 8$

(2) 10個の値の2乗の平均値を  $a$  とすると

$\rightarrow (a - 9^2) = 3$

すなわち  $a = 84$

残りの5個の値の2乗の平均値を  $b$  とすると

$\rightarrow (b - 6^2) = 9$

すなわち  $b = 45$

よって、15個の値の2乗の和は

$a \times 10 + b \times 5 = 84 \times 10 + 45 \times 5 = 1065$

したがって、15個の値の分散は

『2乗の平均値』-『平均値の2乗』  $\rightarrow \frac{1065}{15} - 8^2 = 71 - 64 = 7$

**度数分布から標準偏差を求める**

度数分布から標準偏差を求める場合は、各偏差に対応する度数を掛けたものの平均値を用いれば良い。

**練習)** 下の表は、あるクラス40人に数学の小テストを行った結果である。このクラスの小テストの得点の標準偏差を小数第3位を四捨五入して、小数第2位まで求めよ。ただし、 $\sqrt{3} = 1.732$  とする。

得点 $x$	度数 $f$	$xf$	偏差 $x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$
0	2	0	-4	16	32
1	4	4	-3	9	36
2	7	14	-2	4	28
3	8	24	-1	1	8
4	4	16	0	0	0
5	6	30	1	1	6
6	2	12	2	4	8
7	1	7	3	9	9
8	2	16	4	16	32
9	3	27	5	25	75
10	1	10	6	36	36
計	40 (人)	160			270
平均		$\bar{x} = 4$		分散 $s^2$	6.75
				標準偏差 $s$	2.60

平均や分散の計算は  
度数を掛けたものの  
平均値を計算すること

$$\begin{aligned} \sqrt{\frac{270}{40}} &= \sqrt{\frac{27}{4}} \\ &= \frac{3\sqrt{3}}{2} \\ &= 2.598 \\ &\approx 2.60 \end{aligned}$$

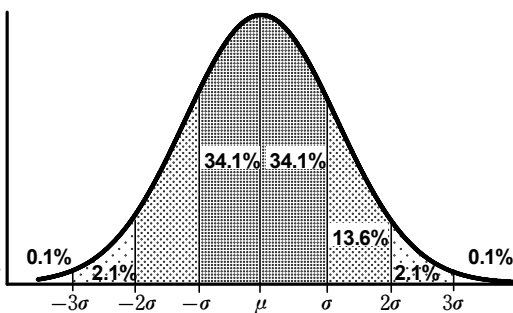
**標準偏差と偏差値**

四分位範囲にはデータの約50%が含まれることとなります。一方、平均値  $\mu$  からのずれが  $\pm$ 標準偏差  $\sigma$  以下の範囲には 68.27%、 $\pm 2 \times$ 標準偏差  $\sigma$  以下だと 95.45%、さらに  $\pm 3 \times$ 標準偏差  $\sigma$  だと 99.73% となります。

このことを用いて模擬試験などでは偏差値として数値化し、全体（母集団）との位置関係を示しています。公式は次のようになり

$$(\text{偏差値}) = \frac{10 \times \{(\text{得点}) - (\text{平均点})\}}{(\text{標準偏差})} + 50$$

「+50」とするので偏差値 50 が集団の中央(平均点)となるのです。ただしあくまでも分布内の数値であり、確率に関わる数値なので目安として捉えましょう。たとえば、2回のテストの点数がそれぞれ、81点、76点であっても平均点や標準偏差によって表のように偏差値が求められます。第1回より第2回のテストの点数は下がっていますが、偏差値は上がっています。偏差値は全体の中で自分がどの位置にいるかを示していますから、この場合は、がっかりしなくてもよいかもしれません。



	得点	平均値	標準偏差	偏差値
第1回	81	60	14	65
第2回	76	60	8	70