

第 40 講 標準偏差

学習目標

- ・ 分散と標準偏差の関係を説明できる。
- ・ 分散の正の平方根として標準偏差を求められる。
- ・ 基本的なデータから平均値、分散、標準偏差を順に求められる。
- ・ 標準偏差が元のデータと同じ単位をもつことを理解し、データの散らばりを比較できる。

分散をそのまま読むと単位が変わる

分散は散らばりを表すが、そのままでは単位が違う

前講では、各データと平均値との差である偏差を 2 乗し、その平均をとったものを分散としました。たとえば、長さを cm で測ったデータを考えます。

偏差の単位は cm ですが、それを 2 乗した偏差の 2 乗の単位は cm^2 になります。

したがって、その平均である分散の単位も cm^2 です。

分散は散らばりの大きさを表せますが、元のデータとは単位が異なります。

そこで、分散の平方根を取り、元のデータと同じ単位に戻した量を使います。

これが**標準偏差**です。

標準偏差

分散の正の平方根を**標準偏差**といいます。

分散を V 、標準偏差を s とすると、

$$s = \sqrt{V}$$

です。

標準偏差は 0 以上の値です。

分散から標準偏差へ

標準偏差を求めるときは、

1. 分散を求める。
2. 分散の正の平方根をとる。

という順に進めます。

分散が 9 なら、標準偏差は 3 です。

正の平方根を使う

分散が 16 のとき、

$$\sqrt{16} = 4$$

なので、標準偏差は 4 です。

−4 は標準偏差ではありません。

標準偏差は散らばりの大きさを表すので、負の値にはなりません。

基本データから標準偏差を求める

平均値から順に計算する

データ

2, 4, 6

の標準偏差を求めます。

平均値は 4 です。

偏差は -2, 0, 2 なので、偏差の 2 乗は 4, 0, 4 です。

したがって、分散は

$$\frac{4 + 0 + 4}{3} = \frac{8}{3}$$

です。

よって、標準偏差は

$$\sqrt{\frac{8}{3}}$$

です。

標準偏差が整数になるとは限りません。

データから標準偏差を求める手順

データから標準偏差を求めるときは、

1. 平均値を求める。
2. 各データの偏差を求める。
3. 偏差を 2 乗する。
4. 偏差の 2 乗の平均をとって分散を求める。
5. 分散の正の平方根をとる。

という順に計算します。

表を使って整理する

偏差と偏差の 2 乗を並べる

データ

1, 3, 5, 7

の標準偏差を求めます。

平均値は 4 です。

データ x	偏差 $x - 4$	偏差の 2 乗 $(x - 4)^2$
1	-3	9
3	-1	1
5	1	1
7	3	9

偏差の 2 乗の合計は 20 です。

したがって、分散は

$$\frac{20}{4} = 5$$

です。
よって、標準偏差は

$$\sqrt{5}$$

です。

平方根をとるのは最後

偏差を求める途中で平方根をとるのではありません。
まず偏差の 2 乗の平均として分散を求め、そのあと分散の平方根をとります。

標準偏差は元のデータと同じ単位になる

単位を元に戻して散らばりを見る

ある長さのデータを cm で表しているとします。
平均との差である偏差の単位は cm です。
偏差を 2 乗すると単位は cm^2 になるので、分散の単位も cm^2 です。
標準偏差では分散の平方根をとるため、単位は再び cm になります。
したがって、標準偏差は元のデータと同じ単位で散らばりを表せます。

標準偏差の単位

元のデータの単位が、

- cm なら標準偏差も cm
- kg なら標準偏差も kg
- 点なら標準偏差も点

となります。
元のデータと同じ単位なので、散らばりの大きさを実際の量として捉えやすくなります。

標準偏差の大きさを散らばりを比べる

大きいほど平均値から広く散らばる

次の 2 つのデータを考えます。
データ A :

4, 5, 6

データ B :

1, 5, 9

どちらも平均値は 5 です。
データ A の分散は $\frac{2}{3}$ なので、標準偏差は

$$\sqrt{\frac{2}{3}}$$

です。
データ B の分散は $\frac{32}{3}$ なので、標準偏差は

$$\sqrt{\frac{32}{3}}$$

です。
データ B の方が標準偏差が大きいので、平均値のまわりにより大きく散らばっていると分かります。

標準偏差と散らばり

平均値のまわりの散らばりを比べるとき、

- ・ 標準偏差が小さいほど、データは平均値の近くに集まっている。
- ・ 標準偏差が大きいほど、データは平均値から広く散らばっている。

と考えます。

平均値も標準偏差も見方が違う

平均値はデータの中心的な位置を表します。

標準偏差は、平均値のまわりにどの程度散らばっているかを表します。

平均値が同じでも標準偏差が異なることがあります。

すべて同じ値なら標準偏差は 0

散らばりがないデータ

データ

6, 6, 6, 6

の平均値は 6 です。

すべての偏差が 0 なので、分散は 0 です。

したがって、標準偏差も

$$\sqrt{0} = 0$$

です。

標準偏差が 0 ということは、すべてのデータが平均値と同じで、散らばりがないことを表します。

例題

例題 1

次のデータの分散と標準偏差を求めなさい。

2, 4, 4, 6

例題 1 の解答・解説

平均値は 4 です。

偏差は $-2, 0, 0, 2$ なので、偏差の 2 乗は $4, 0, 0, 4$ です。

分散は

$$\frac{4 + 0 + 0 + 4}{4} = 2$$

です。

したがって、標準偏差は

$$\sqrt{2}$$

です。

例題 2

次のデータの標準偏差を求めなさい。

1, 1, 5, 5

例題 2 の解答・解説

平均値は

$$\frac{1 + 1 + 5 + 5}{4} = 3$$

です。

偏差は $-2, -2, 2, 2$ です。

偏差の 2 乗はすべて 4 なので、分散は 4 です。

したがって、標準偏差は

$$\sqrt{4} = 2$$

です。

例題 3

次の 2 つのデータは、ともに平均値が 10 です。

データ A の分散は 4、データ B の分散は 25 です。

それぞれの標準偏差を求め、どちらの方が平均値のまわりに大きく散らばっているか答えなさい。

例題 3 の解答・解説

データ A の標準偏差は

$$\sqrt{4} = 2$$

です。

データ B の標準偏差は

$$\sqrt{25} = 5$$

です。

標準偏差はデータ B の方が大きいので、データ B の方が平均値のまわりに大きく散らばっています。

演習問題

練習問題 1

分散が 9 であるデータの標準偏差を求めなさい。

練習問題 2

分散が 7 であるデータの標準偏差を求めなさい。

練習問題 3

次のデータの分散と標準偏差を求めなさい。

3, 5, 7

練習問題 4

次のデータの標準偏差を求めなさい。

2, 2, 6, 6

練習問題 5

次のデータの標準偏差を求めなさい。

4, 4, 4, 4, 4

練習問題 6

次のデータの分散と標準偏差を求めなさい。

1, 3, 5, 7, 9

練習問題 7

2 つのクラス A、B のテスト得点について、平均値はいずれも 70 点でした。

クラス A の分散は 16、クラス B の分散は 36 です。

(1) それぞれの標準偏差を求めなさい。

(2) どちらのクラスの方が、得点が平均値のまわりに大きく散らばっていると考えられるか答えなさい。

(3) 標準偏差の単位を答えなさい。

練習問題 8

ある生徒は、長さを cm で測ったデータの分散が 25 であることから、「標準偏差は 25 cm である」と答えました。

この答えの誤りを説明し、正しい標準偏差を求めなさい。

演習問題の解答・解説**練習問題 1 の解答・解説**

標準偏差は分散の正の平方根です。

したがって、

$$\sqrt{9} = 3$$

です。

練習問題 2 の解答・解説

標準偏差は分散の正の平方根なので、

$$\sqrt{7}$$

です。

練習問題 3 の解答・解説

平均値は 5 です。

偏差は $-2, 0, 2$ なので、偏差の 2 乗は $4, 0, 4$ です。

分散は

$$\frac{4 + 0 + 4}{3} = \frac{8}{3}$$

です。

したがって、標準偏差は

$$\sqrt{\frac{8}{3}}$$

です。

練習問題 4 の解答・解説

平均値は

$$\frac{2 + 2 + 6 + 6}{4} = 4$$

です。

偏差は $-2, -2, 2, 2$ なので、偏差の 2 乗はすべて 4 です。

分散は 4 です。

したがって、標準偏差は 2 です。

練習問題 5 の解答・解説

平均値は 4 です。

すべての偏差が 0 なので、分散は 0 です。

したがって、標準偏差も 0 です。

練習問題 6 の解答・解説

平均値は

$$\frac{1 + 3 + 5 + 7 + 9}{5} = 5$$

です。

偏差は $-4, -2, 0, 2, 4$ です。

偏差の 2 乗は $16, 4, 0, 4, 16$ なので、分散は

$$\frac{16 + 4 + 0 + 4 + 16}{5} = 8$$

です。

したがって、標準偏差は

$$\sqrt{8} = 2\sqrt{2}$$

です。

練習問題 7 の解答・解説

(1) クラス A の標準偏差は

$$\sqrt{16} = 4$$

です。

クラス B の標準偏差は

$$\sqrt{36} = 6$$

です。

(2) 標準偏差が大きいクラス B の方が、得点が平均値のまわりに大きく散らばっています。

(3) 元のデータがテスト得点なので、標準偏差の単位も点です。

練習問題 8 の解答・解説

標準偏差は分散そのものではなく、分散の正の平方根です。
分散が 25 なので、

$$\sqrt{25} = 5$$

です。

したがって、標準偏差は 5 cm です。

標準偏差は元のデータと同じ単位になります。