

平均偏差（絶対平均偏差）からの分散、標準偏差の導入

札幌月寒高等学校 教諭 正田 隆之

データの散らばりを数値化する際、平均偏差で表すと感覚的に分かりやすいので、分散、標準偏差への橋渡しとして使っています。その導入事例を紹介します。授業では教科書の例に沿った形にアレンジしています。誤り、理解不足があるかもしれませんがその際にご指摘ください。

平均値 データ全体を**大雑把に1つの数値**で表現する

テニス部の男子10人の身長が以下の表のとおりであった。

| | | | | | | | | | | |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 部員名 | a | b | c | d | e | f | g | H | i | j |
| 身長(cm) | 159 | 178 | 155 | 185 | 163 | 183 | 158 | 180 | 162 | 177 |

問1 部員10人の身長の平均値を求めよ。 → 170 cm

個々の身長は様々であるが、「部員の身長は170 cm程度」と大雑把に判断できる。

これを根拠にして、部員全員に一律同じLサイズのTシャツを配布して着用させることにしたが、実際にはTシャツが大きすぎたり、小さすぎたりして心地よくない部員がいる可能性がある。それを判断するための指標について考える。

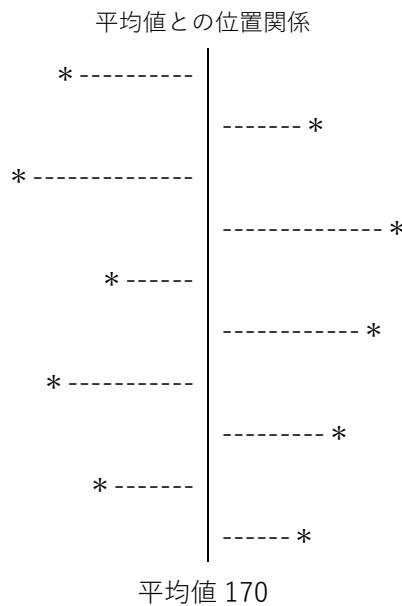
偏差 平均値と比べてどれだけ大きい、どれだけ小さい？

問2 部員それぞれの身長の「偏差」を求めよ。

例 部員aの身長の「偏差」は $159 - 170$ (平均値) = -11

→ 部員aは平均より11 cm低い

| 部員 | 身長 | 偏差 |
|----|-----|-----|
| a | 159 | -11 |
| b | 178 | 8 |
| c | 155 | -15 |
| d | 185 | 15 |
| e | 163 | -7 |
| f | 183 | 13 |
| g | 158 | -12 |
| h | 180 | 10 |
| i | 162 | -8 |
| j | 177 | 7 |



身長が平均値から10 cm以上離れている部員が6人もいます。平均値だけで、全員の身長を大雑把に判断していたが、実は、平均値近くの人ばかりではなく意外と離れた値まで散らばっていることがわかる。Tシャツがフィットしない部員がけっこういそうである。ここでは個々の数値あるいは、それをグラフで表して全体を眺めて判断したが、1つの数値で表現できないだろうか。

平均偏差 散らばり方を**大雑把に1つの数値**で表現する

それぞれの部員の身長「平均値からの距離」に着目する。

例 部員 *a* の身長「平均値からの距離」は $|159 - 170| = |-11| = 11$
つまり、平均値との距離は「偏差の絶対値」で表せる。

問3 部員それぞれの身長「平均値からの距離」を求めよ。また、それらの値の平均値を求めよ。

| 部員名 | <i>a</i> | <i>b</i> | <i>c</i> | <i>d</i> | <i>e</i> | <i>f</i> | <i>g</i> | <i>H</i> | <i>i</i> | <i>j</i> | 平均値 |
|------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|------|
| 身長(cm) | 159 | 178 | 155 | 185 | 163 | 183 | 158 | 180 | 162 | 177 | 170 |
| 偏差 | -11 | 8 | -15 | 15 | -7 | 13 | -12 | 10 | -8 | 7 | (0) |
| 平均値からの距離(偏差の絶対値) | 11 | 8 | 15 | 15 | 7 | 13 | 12 | 10 | 8 | 7 | 10.6 |

「偏差の絶対値」を平均した値を「平均偏差(または絶対平均偏差)」という。すなわち、この部員たちの身長「平均偏差は10.6 cm」である。実際の個々の身長「平均値からの距離」は様々であるが「部員の身長は平均値から10.6 cm前後離れたところまで散らばっている」と大雑把に判断できる。

データの散らばりを「平均偏差」で表現すると感覚的に分かり易いが、様々な計算をする際に絶対値が邪魔になる(積分の計算が困難)らしく、「平均偏差」が使われることはあまりない。実際には散らばり方をどのように表現しているのだろうか。

分散 データの散らばりを**大きさに表現**する

問4 部員それぞれの「身長の偏差を2乗した値」を求めよ。また、それらの平均値を求めよ。

| 部員名 | <i>a</i> | <i>b</i> | <i>c</i> | <i>d</i> | <i>e</i> | <i>f</i> | <i>g</i> | <i>H</i> | <i>i</i> | <i>j</i> | 平均値 |
|--------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----|
| 身長(cm) | 159 | 178 | 155 | 185 | 163 | 183 | 158 | 180 | 162 | 177 | 170 |
| 偏差 | -11 | 8 | -15 | 15 | -7 | 13 | -12 | 10 | -8 | 7 | (0) |
| 偏差の2乗 | 121 | 64 | 225 | 225 | 19 | 169 | 144 | 100 | 64 | 49 | 121 |

「偏差の2乗」は、「平均値からどれだけ離れているか」をかなり大きさに表現したものと考えられる。この「偏差の2乗」を平均した値を「分散」という。すなわち、この部員たちの身長の分散は121である。これにより「部員の身長は大きさに表現して平均値170 cmからの距離の2乗の値が121前後まで散らばっている」ことになる。しかし、かなり大きな値のため感覚的に分かりづらい。

標準偏差 **大きさにってしまった分を中和させて自然な表現に近づける**

「分散」の計算の際、「個々の偏差を2乗」しているから、逆に計算結果「分散」の「平方根」を計算すれば、平均偏差と同じような感覚で判断できるのではないだろうか? 今回の例の「分散の平方根」と「平均偏差」とを比較すると次のようになる。

$$(\text{分散の平方根}) \sqrt{121} = 11 \approx 10.6 (\text{平均偏差}) \dots\dots \text{ほぼ等しい}$$

ここまで近い値になるとは限らないが、分散にくらべれば、かなり平均偏差に感覚的に近くなる。この「分散の平方根」を「標準偏差」という。すなわち、部員たちの身長の「標準偏差は11 cm」である。これにより、「部員の身長は平均値170 cmから11 cm前後離れたところまで散らばっている」と大雑把に判断できる。