数学Ⅰ　第５章　データの分析　No.4

学習のねらい

分散と平均の関係について理解しよう！

変量の変換について理解しよう！

1. 分散と平均の関係

分散s^2を変形してみよう。

$$s^2 = \frac{1}{n}\{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \cdots + (x_n - \overline{x})^2\}$$

$$= \frac{1}{n}\{(x_1^2 + x_2^2 + \cdots x_n^2) - 2\overline{x}(x_1 + x_2 + \cdots + x_n) + n(\overline{x})^2\}$$

$$= \frac{1}{n}(x_1^2 + x_2^2 + \cdots x_n^2) - 2\overline{x} \cdot \frac{1}{n}(x_1 + x_2 + \cdots + x_n) + (\overline{x})^2$$

$$= \overline{x^2} - 2\overline{x} \cdot \overline{x} + (\overline{x})^2$$

$$= \overline{x^2} - (\overline{x})^2$$　（$\overline{x^2}$とは、x_1^2、x_2^2、$\cdots x_n^2$の平均のことである。）

よって、以下の公式が得られる。

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

実際の計算では、こちらの公式を使うと楽になることが多い。

2. 変量の変換

テストとかで、全員の点数を0.8倍するとか、全員の点数に+10点するっていう場面に出会ったことはないかな？

ここでは、データの各値に一斉に同じ数を加えたり、一斉に同じ数を掛けたりする時、平均値、分散、標準偏差がどのようになるかを考えよう。

変量xについてのデータがn個の値x_1、x_2、$\cdots$、x_nであるとして、xのデータの平均値を$\overline{x}$、分散をs_x^2、標準偏差をs_xとする。そして、a、bを定数として、式$y = ax + b$で新しい変量yを作る。このyの平均値を$\overline{y}$、分散をs_y^2、標準偏差をs_yとすると以下が成り立つ。（証明は◇問題で。）

$$\overline{y} = a\overline{x} + b \qquad s_y^2 = a^2 s_x^2 \qquad s_y = |a| s_x$$

つまり、分散、標準偏差には$+b$の部分は影響しないということ。考えてみたら当たり前で、分散と標準偏差は、散らばり具合だから、$+b$したところで、散らばり具合は変わらない。

数学Ⅰあと少しで終わり！！！！

最後まで駆け抜けよう！！

◊問題

1.　25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4、分散は14、残りの15個の値の平均値は9、分散は19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。

2.　a、bは定数とする。変量xのデータから$y = ax + b$によって新しい変量yのデータが得られるとき、x、yのデータの平均値を$\overline{x}$、$\overline{y}$、分散をs_x^2、s_y^2、標準偏差をs_x、s_yとするとき(1)～(3)の式が成り立つことを示せ。

（1）$\overline{y} = a\overline{x} + b$　（2）$s_y^2 = a^2 s_x^2$　（3）$s_y = |a| s_x$