

それでは、もう少し大きなデータを使って、分散  $S^2$  と標準偏差  $S$  を求めてみよう。ここでは、さらに、仮平均の利用の仕方も教えよう。

**練習問題 38**

分散と標準偏差

CHECK 1

CHECK 2

CHECK 3

12 人の子供が、あるゲームを行った結果、次の得点データが得られた。

13, 11, 15, 18, 11, 14, 17, 12, 11, 16, 20, 10

このデータの分散  $S^2$  と標準偏差  $S$  を求めよ。

まず、 $x_1 = 13, x_2 = 11, \dots, x_{12} = 10$  において、公式通りに、平均値  $m$ 、分散  $S^2$ 、標準偏差  $S$  を求めていけばいいんだね。その際に、表を利用すると、ミスを出さずに計算できると思う。

12 個のデータを、

$x_1 = 13, x_2 = 11, x_3 = 15, x_4 = 18, x_5 = 11, x_6 = 14, x_7 = 17, x_8 = 12,$

$x_9 = 11, x_{10} = 16, x_{11} = 20, x_{12} = 10,$

とにおいて、この平均値  $m$  を求めると、

$$\begin{aligned}
 m &= \frac{1}{12} (x_1 + x_2 + x_3 + \dots + x_{12}) \\
 &= \frac{1}{12} (\underbrace{13+11}_{24} + \underbrace{15+18}_{39} + \underbrace{11+14}_{25} + \underbrace{17+12}_{29} + \underbrace{11+16}_{27} + \underbrace{20+10}_{30}) \\
 &= \frac{168}{12} = \frac{42}{3} = 14 \text{ となる。}
 \end{aligned}$$

**参考**

12 個のデータは、10 から 20 の値を取っているのだから、勘でまず仮平均  $m'$  として、 $m' = 15$  としよう。すると、本当の平均値  $m$  は、この  $m'$  に、各データの  $m'$  との偏差の平均を加えたものになる。

各データと  $m'$  との偏差の平均

$$\begin{aligned}
 m &= \underbrace{15}_{\text{仮平均 } m'} + \frac{1}{12} (\underbrace{-2}_{x_1 - m'} - 4_{x_2 - m'} + 0 + 3 - 4 - 1 + 2 - 3 - 4 + 1 + \underbrace{5}_{x_{11} - m'} + \underbrace{5}_{x_{12} - m'}) \\
 &= 15 + \frac{1}{12} \times (-12) = 15 - 1 = 14 \text{ となる。}
 \end{aligned}$$

これだと、扱う数字が小さくなるので、計算がしやすくなるんだね。

したがって、偏差平方の平均が分散  $S^2$  より、

偏差平方の和をデータの個数 12 で割ったもの

$$\begin{aligned}
 S^2 &= \frac{1}{12} \{ \underbrace{(x_1 - m)^2}_{(13-14)^2} + \underbrace{(x_2 - m)^2}_{(11-14)^2} + \underbrace{(x_3 - m)^2}_{(15-14)^2} + \\
 &\quad \cdots + \underbrace{(x_{11} - m)^2}_{(20-14)^2} + \underbrace{(x_{12} - m)^2}_{(10-14)^2} \} \\
 &= \frac{1}{12} (1 + 9 + 1 + 16 + 9 + 0 + \cdots + 16) \\
 &= \frac{114}{12} = \frac{19}{2} = 9.5 \text{ となる。}
 \end{aligned}$$

よって、この正の平方根が標準偏差  $S$  となるので、

$$S = \sqrt{S^2} = \sqrt{\frac{19}{2}} = \frac{\sqrt{38}}{2}$$

≒ 3.082 となって、答えだ。

表

データ No.	データ	偏差 $x_k - m$	偏差平方 $(x_k - m)^2$
$x_1$	13	-1	1
$x_2$	11	-3	9
$x_3$	15	1	1
$x_4$	18	4	16
$x_5$	11	-3	9
$x_6$	14	0	0
$x_7$	17	3	9
$x_8$	12	-2	4
$x_9$	11	-3	9
$x_{10}$	16	2	4
$x_{11}$	20	6	36
$x_{12}$	10	-4	16
合計	168	0	114
平均	14		9.5

平均値  $m$

分散  $S^2$

以上で、今日の講義は終了です。平均値やメジアン(中央値)やモード(最頻値)、それに第1, 第2, 第3四分位数, …など、様々な統計用語が出てきたので、まず、これらの用語の意味をヨ〜ク頭の中で整理しておくことだね。

そして、後半は平均値や分散や標準偏差の計算練習をやったんだね。これらの計算は試験では頻出なので、表も利用しながら正確に迅速に結果が出せるように、何度でも自分で納得がいくまで練習しておこう。

では、次回は、もう少し複雑な2変数データの分析について解説するけれど、また超分かりやすく解説するから、すべて理解できるはずだ。みんな楽しみに待っていてくれ。それでは、次回の講義でまた会おう。

みんな元気で…。さようなら…。