

平均・分散・標準偏差

意味がわかる統計学

第1章 p.26~43

60216013

池内 未来

目次

- 平均・分散・標準偏差
- 平均・分散・標準偏差(度数分布表から)
- 平均・分散・標準偏差(相対度数分布表から)
- 平均・分散・標準偏差(ヒストグラムから)
- まとめ

平均・分散・標準偏差

平均・分散・標準偏差の公式

資料の個数がN個, 資料の変量が $x_1, x_2, \dots, x_N$ のとき
資料の平均 $\bar{x}$, 分散 s^2 , 標準偏差 s はそれぞれ次のようになる

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_N}{N}$$

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \dots + (x_N - \bar{x})^2}{N}$$

$$s = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \dots + (x_N - \bar{x})^2}{N}}$$

例題

43、 47、 52、 52、 54、 61、 67、 67、 68、 69、
70、 71、 71、 73、 76、 78、 82、 84、 84、 91

クラスの人数は20人で、点数は上記のようになっている
平均・分散・標準偏差を求める

例題

43、 47、 52、 52、 54、 61、 67、 67、 68、 69、
70、 71、 71、 73、 76、 78、 82、 84、 84、 91

- クラス20人の合計点は1360点なので平均は

$$\bar{x} = \frac{1360}{20} = 68$$

- 分散は

$$s^2 = \frac{(-25)^2 + (-21)^2 + \dots + 16^2 + 23^2}{20} = \frac{3274}{20} = 163.7$$

- 標準偏差は

$$s = \sqrt{163.7} \approx 12.79$$

平均・分散・標準偏差 (度数分布表から)

平均・分散・標準偏差の公式

度数分布表が

階級値	度数
x_1	f_1
x_2	f_2
$\vdots$	$\vdots$
x_n	f_n
合計	N

であるとき

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_n f_n}{N}$$

$$s^2 = \frac{(x_1 - \bar{x})^2 f_1 + (x_2 - \bar{x})^2 f_2 + \dots + (x_n - \bar{x})^2 f_n}{N}$$

$$s = \sqrt{\frac{(x_1 - \bar{x})^2 f_1 + (x_2 - \bar{x})^2 f_2 + \dots + (x_n - \bar{x})^2 f_n}{N}}$$

例題

階級	階級値	度数
40 以上 50 未満	45	2
50 以上 60 未満	55	3
60 以上 70 未満	65	5
70 以上 80 未満	75	6
80 以上 90 未満	85	3
90 以上 100 以下	95	1

上の度数分布表から,平均・分散・標準偏差を求める

例題

階級	階級値	度数
40 以上 50 未満	45	2
50 以上 60 未満	55	3
60 以上 70 未満	65	5
70 以上 80 未満	75	6
80 以上 90 未満	85	3
90 以上 100 以下	95	1

- 階級値の点数の人が度数の人数分いると捉える

$$\bar{x} = \frac{45 \times 2 + 55 \times 3 + 65 \times 5 + 75 \times 6 + 85 \times 3 + 95 \times 1}{20} = \frac{1380}{20} = 69$$

$$s^2 = \frac{(-24)^2 \times 2 + (-14)^2 \times 3 + (-4)^2 \times 5 + 6^2 \times 6 + 16^2 \times 3 + 26^2 \times 1}{20} = 174$$

$$s = \sqrt{174} \approx 13.19$$

平均・分散・標準偏差 (相対度数分布表から)

平均・分散・標準偏差の公式

相対度数分布が

階級値	度数	相対度数
x_1	f_1	$\frac{f_1}{N} = g_1$
x_2	f_2	$\frac{f_2}{N} = g_2$
$\vdots$	$\vdots$	$\vdots$
x_n	f_n	$\frac{f_n}{N} = g_n$

であるとき

$$\bar{x} = x_1g_1 + x_2g_2 + \dots x_ng_n$$

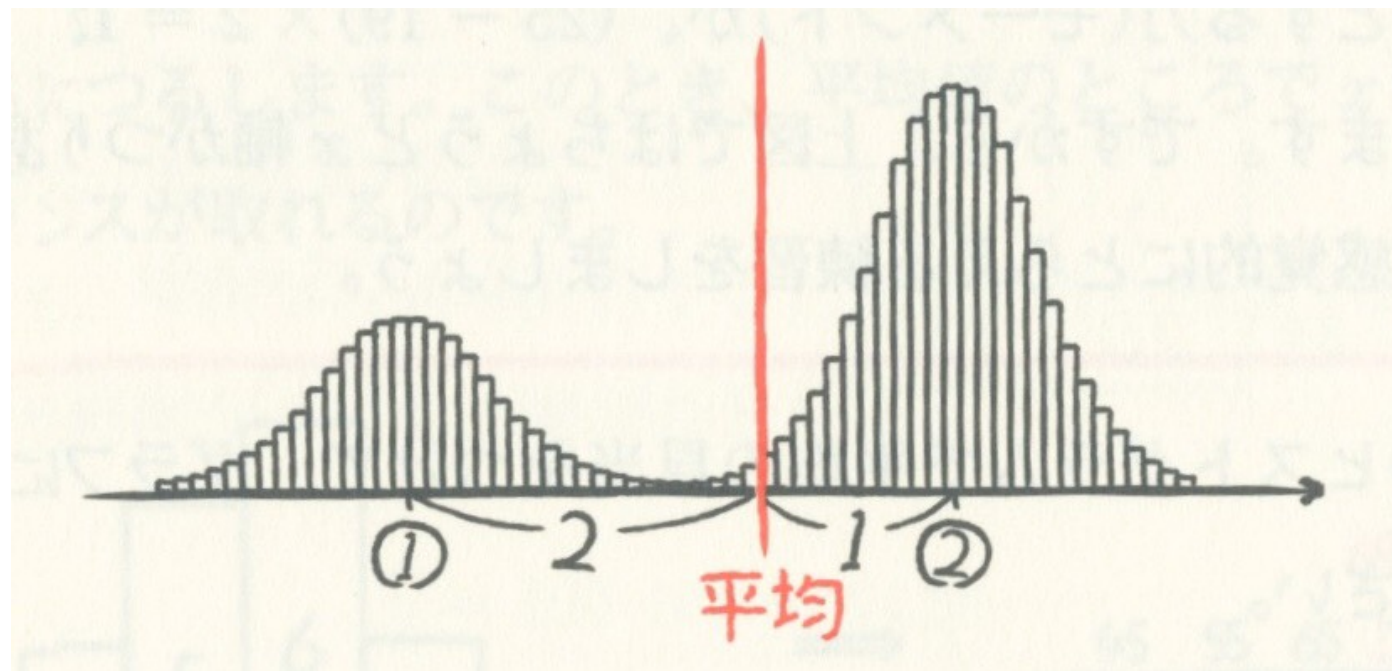
$$s^2 = (x_1 - \bar{x})^2g_1 + (x_2 - \bar{x})^2g_2 + \dots + (x_n - \bar{x})^2g_n$$

$$s = \sqrt{(x_1 - \bar{x})^2g_1 + (x_2 - \bar{x})^2g_2 + \dots + (x_n - \bar{x})^2g_n}$$

平均・分散・標準偏差 (ヒストグラムから)

平均

ヒストグラムをx軸の一点で支えるとき
バランスを取ることができる点



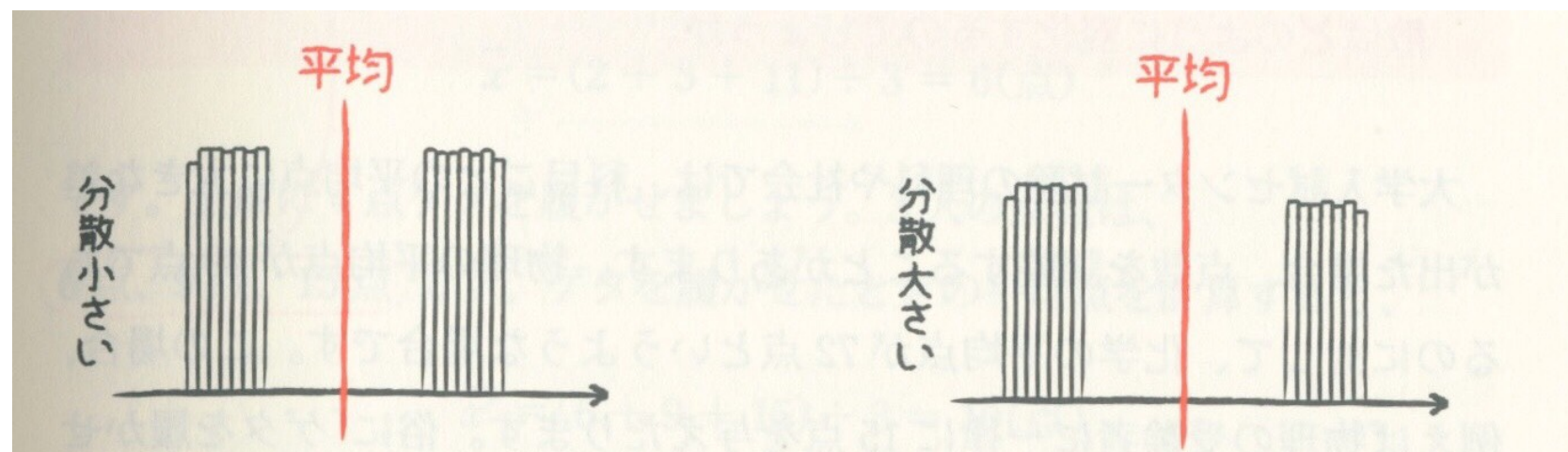
分散・標準偏差

分散：資料の散らばり度合いを表したものの

→分散が大きいと散らばり方が大きい

→分散が小さいと散らばり方が小さい

標準偏差：分散の平方根であるので同様になる



まとめ

まとめ

- 平均：資料の変量の総和を資料の個数で割ったもの
- 分散・標準偏差：資料の散らばり度合いを表したもの
- ヒストグラム：度数分布を示すグラフ