組　　番　名前

1 次のデータは，ある生徒 20 人が 1 年間に図書室から借りた本の冊数である。

17	12	10	10	12	5	7	10	7	9	
9	15	13	1	4	6	14	1	9	8	(冊)

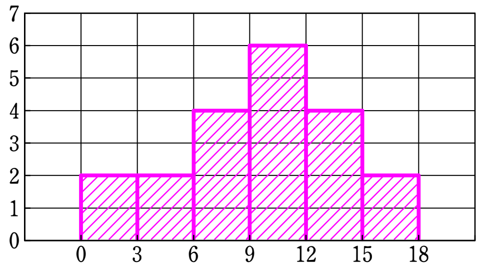
本の冊数	1	1	4	5	6	7	7	8	9	9	9	10	10	10	12	12	13	14	15	17

(1) **階級の幅**を 3 冊として，**度数分布表**を作れ。ただし，**階級**は 0 冊から区切り始めるものとする。

本の冊数の階級 (冊)	度数
0 以上 3 未満	2
3 ～ 6	2
6 ～ 9	4
9 ～ 12	6
12 ～ 15	4
15 ～ 18	2
計	20

(2) もっとも度数が大きい階級の**階級値**は 10.5 である。

(2) (1) で作った度数分布表をもとにして，**ヒストグラム**をかけ。



2 次の 5 個のデータ　3，2，5，1，3　において，各問いに答えよ。

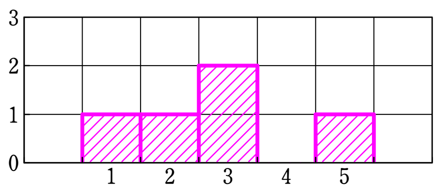
(1) 小さい順に並べよ。

データ	1	2	3	3	5

(2) 度数分布表の空欄を補え。

階級	度数
1	1
2	1
3	2
4	0
5	1
合計	5

(3) ヒストグラムをかけ。



☆ (4)～(13)は小数点以下第1位まで答えよ。必要ならば小数点以下第2位を四捨五入すること。

(4) 平均値は 2 . 8 である。

(5) 中央値（メジアン）は 3 . 0 である。

(6) 最頻値（モード）は 3 . 0 である。

(7) 最大値は 5 . 0 である。

(8) 最小値は 1 . 0 である。

(9) 第 1 四分位数は 1 . 5 である。

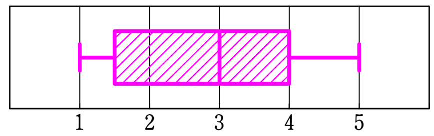
(10) 第 2 四分位数は 3 . 0 である。

(11) 第 3 四分位数は 4 . 0 である。

(12) 四分位範囲は 2 . 5 である。

(13) 四分位偏差は 1 . 3 である。

(14) 箱ひげ図をかけ。（ただし，平均値は記入しなくてよい。）



3 次の 9 個のデータ　5，3，7，4，5，1，8，5，7　において，各問いに答えよ。（解答の様式は下記点線枠内を参考にせよ。）

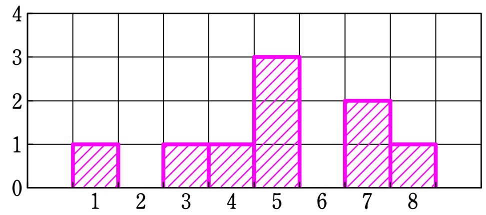
(1) 小さい順に並べよ。

データ	1	3	4	5	5	5	7	7	8

(2) 度数分布表の空欄を補え。

階級	度数
1	1
2	0
3	1
4	1
5	3
6	0
7	2
8	1
計	9

(3) ヒストグラムをかけ。



☆ (4)～(13)は小数点以下第1位まで答えよ。必要ならば小数点以下第2位を四捨五入すること。

(4) 平均値は 5 . 0 である。

(5) 中央値（メジアン）は 5 . 0 である。

(6) 最頻値（モード）は 5 . 0 である。

(7) 最大値は 8 . 0 である。

(8) 最小値は 1 . 0 である。

(9) 第 1 四分位数は 3 . 5 である。

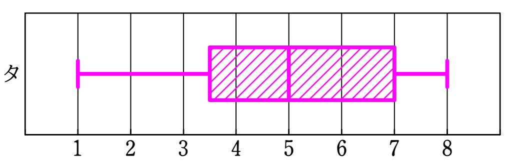
(10) 第 2 四分位数は 5 . 0 である。

(11) 第 3 四分位数は 7 . 0 である。

(12) 四分位範囲は 3 . 5 である。

(13) 四分位偏差は 1 . 8 である。

(14) 箱ひげ図をかけ。（ただし，平均値は記入しなくてよい。）



新編数学A　P.170～182

組　番　名前

1 [714新編 数学Ⅰ 本文ページ182] E　外れ値

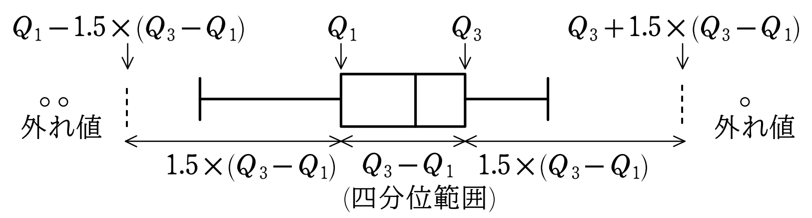
データの中に，他の値から極端に離れた値が含まれることがある。そのような値を**外れ値**という。

外れ値の基準は複数あるが，たとえば，次のような値を外れ値とする。

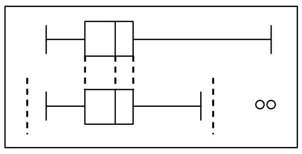
（第1四分位数−1.5×四分位範囲）以下の値　$\leftarrow \{Q_1-1.5\times(Q_3-Q_1)\}$ 以下

（第3四分位数+1.5×四分位範囲）以上の値　$\leftarrow \{Q_3+1.5\times(Q_3-Q_1)\}$ 以上

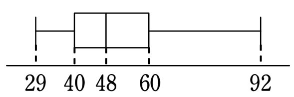
外れ値がある場合，次の図のような箱ひげ図が用いられることがある。
外れ値は∘で示している。また，箱ひげ図の左右のひげは，データから外れ値を除いたときの最小値または最大値まで引いている。



注意　外れ値を∘で示す箱ひげ図をかく場合でも，四分位数は外れ値を除かないすべての値のデータの四分位数であり，その値にもとづいて箱をかく。



練習9　右の図はあるデータの箱ひげ図である。このデータの最大値92，最小値29は外れ値であるかを，四分位範囲を利用して調べよ。

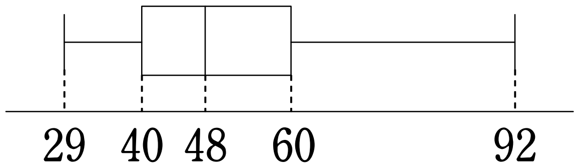


2 に取り出してあります。

外れ値は，測定ミスや入力ミスによる異常な値であることもある。一方で，外れ値の背景を探ることにより，問題発見があったり，問題解決の手がかりが得られたりすることもある。たとえば，販売員の販売成績を調べたとき，並外れて成績が良い販売員がいたら，その販売員の工夫を探ることで，全体の販売成績を上げる対策を見いだせる可能性がある。

2 [714新編 数学Ⅰ 練習9]

次の図はあるデータの箱ひげ図である。
このデータの最大値92，最小値29は外れ値であるかを，四分位範囲を利用して調べよ。



解答　最大値92は外れ値である。　　最小値29は外れ値ではない。

解説　このデータの四分位範囲は　$60-40=20$

$60+1.5\times20=90$ から，90以上の値が外れ値である。

よって，最大値92は外れ値である。

$40-1.5\times20=10$ から，10以下の値が外れ値である。

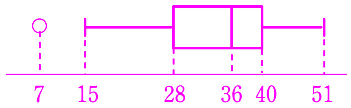
よって，最小値29は外れ値ではない。

3 [3TRIAL数学Ⅰ 問題282]

次のデータについて，外れ値があるかどうかを，四分位範囲を利用して調べよ。また，箱ひげ図をかけ。ただし，外れ値がある場合は○で示せ。

7，15，28，30，35，37，38，40，42，51

解答　外れ値は7である



このデータの最小値，第1四分位数，中央値，第3四分位数，最大値は，順に

$7，28，\frac{35+37}{2}=36，40，51$

また，このデータの四分位範囲は　$40-28=12$

$40+1.5\times12=58$，$28-1.5\times12=10$ から，58以上の値，10以下の値が外れ値である。

よって，外れ値は7である。

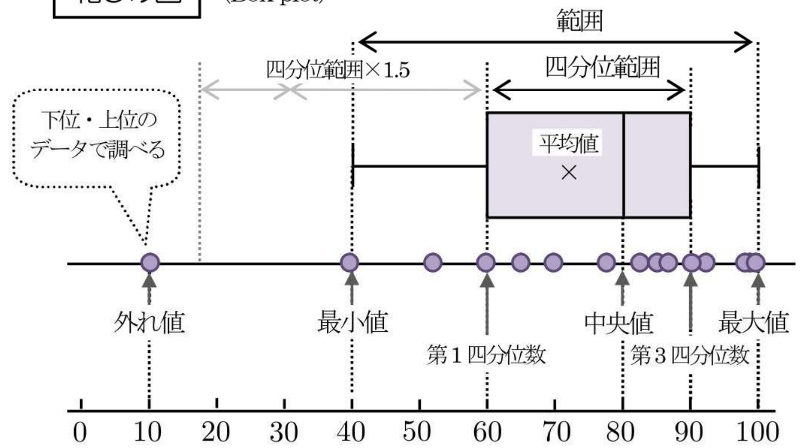
また，外れ値を○で示して箱ひげ図をかくと，上の図のようになる。

4 【数学のいずみ】中村文則　＆　白チャート P.260

変量・データ

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	平均
変量	10	40	52	60	64	70	78	82	85	86	90	92	99	100	72.0
区分	下位データ							上位データ							

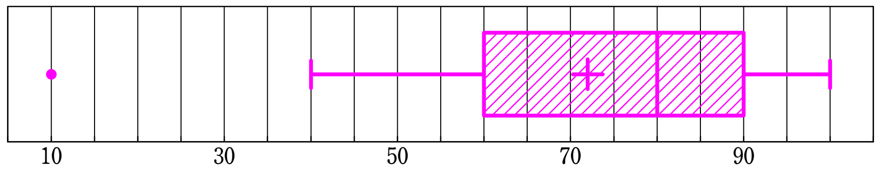
箱ひげ図　(Box plot)



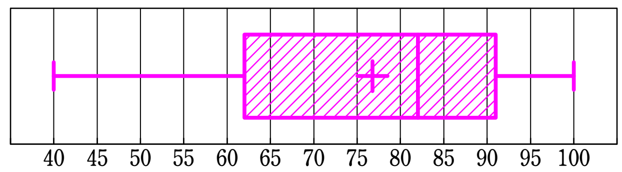
変量に基づいて箱ひげ図を描いてみた。

変量「10」が「外れ値」であることを確認してみよう。

【外れ値を含む】



【外れ値を除外】



【代表値の値の変化】

列1	変量を元に計算	列1	変量を元に計算
データの大きさ	14	データの大きさ	13
平均値	72.00	平均値	76.77
中央値(メジアン)	80.00	中央値(メジアン)	82.00
最頻値(モード)	6件以上あります	最頻値(モード)	6件以上あります
最大値	100.00	最大値	100.00
最小値	10.00	最小値	40.00
第1四分位数	60.00	第1四分位数	62.00
第2四分位数	80.00	第2四分位数	82.00
第3四分位数	90.00	第3四分位数	91.00
四分位範囲	30.00	四分位範囲	29.00
四分位偏差	15.00	四分位偏差	14.50

【白チャート P.260～261】

■ **外れ値**

データの中に，他の値から極端に離れた値が含まれることがある。そのような値を**外れ値**という。

外れ値の基準として次のようなものがある。

{(第1四分位数)−1.5×(四分位範囲)} 以下の値

{(第3四分位数)+1.5×(四分位範囲)} 以上の値

外れ値の影響について，例えば，次のようなものがある。

① 平均値は中央値より外れ値の影響を受けやすい。

② 四分位範囲は外れ値の影響を受けにくい。

1

分散と標準偏差

変量 x のデータの値が x_1, x_2, ……, x_n で，その平均値が $\overline{x}$ のとき

分散　s^2＝偏差の2乗の平均

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

標準偏差　$s=\sqrt{\text{分散}}$

例　次のデータは，5人の生徒にテストを行った結果である。このデータの分散，標準偏差を求めよ。

3，4，6，2，5

解答　この変量を x とすると，データの平均値 $\overline{x}$ は

$$\overline{x}=\frac{1}{5}(3+4+6+2+5)=\frac{\boxed{20}}{5}=\boxed{4}$$ である。

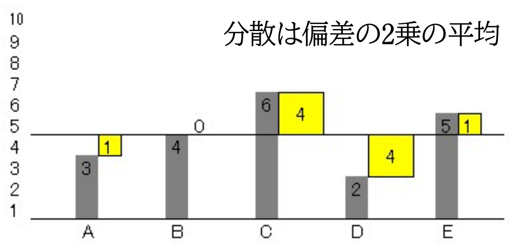
よって，$x-\overline{x}$，$(x-\overline{x})^2$ は次のようになる。

x	3	4	6	2	5	計
$x-\overline{x}$	−1	0	2	−2	1	計　0
$(x-\overline{x})^2$	1	0	4	4	1	計 10

よって，分散 s^2 は　$s^2=\dfrac{1}{\boxed{5}}\times\boxed{10}=\boxed{2}$　答

標準偏差 s は　$s=\sqrt{\boxed{2}}$　答

【分散のイメージ図】



別解　**分散の求め方**　分散を表す式を変形してみよう。

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

(x のデータの分散)＝(x^2 のデータの平均値)－(x のデータの平均値)^2

上の例と同じデータで計算してみよう。

x	3	4	6	2	5	計 20
x^2	9	16	36	4	25	計 90

よって，分散 s^2 は　$s^2=\dfrac{\boxed{90}}{\boxed{5}}-\boxed{4}^2$

$=\boxed{18}-\boxed{16}=\boxed{2}$　答

2　次のデータの分散，標準偏差を求めよ。表の空欄と□を埋めよ。

10人のテストの得点 x が下の表で与えられている。ただし，平均値 $\overline{x}$ は，$\overline{x}=\dfrac{1}{10}\times60=6$ である。x の単位は点である。

x	6	7	4	8	9	4	3	8	6	5	計 60
$x-\overline{x}$	0	1	−2	2	3	−2	−3	2	0	−1	計　0
$(x-\overline{x})^2$	0	1	4	4	9	4	9	4	0	1	計 36

よって，分散 s^2 は　$s^2=\dfrac{1}{\boxed{10}}\times\boxed{36}=\boxed{3.6}$　答

標準偏差 s は　$\sqrt{10}=3.2$ として小数第2位を四捨五入すると

$$s=\sqrt{\boxed{3.6}}=\frac{\boxed{6}}{\sqrt{10}}\fallingdotseq\boxed{1.9}$$（点）

別解 の方法でデータの分散を求めると

x	6	7	4	8	9	4	3	8	6	5	計 60
x^2	36	49	16	64	81	16	9	64	36	25	計 396

$\overline{x}=\dfrac{1}{10}\times60=6$，$\overline{x^2}=\dfrac{1}{10}\times396=39.6$ である。

よって，分散 s^2 は

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

$=\dfrac{\boxed{396}}{\boxed{10}}-\boxed{6}^2=\boxed{39.6}-\boxed{36}=\boxed{3.6}$　答

3　右の表は，ある製品を成型できる2台の工作機械X，Yの1時間あたりのそれぞれの不良品の数 x，y を5時間にわたって調べたものである。(単位は個)

x	5	4	8	12	6
y	6	9	8	5	7

(1)　x，y のデータの平均値，分散，標準偏差をそれぞれ求めよ。ただし，小数第2位を四捨五入せよ。

解答　(1)　x，y のデータの平均値をそれぞれ $\overline{x}$，$\overline{y}$ とすると

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	x^2	y^2
5	6	−2	−1	4	1	25	36
4	9	−3	2	9	4	16	81
8	8	1	1	1	1	64	64
12	5	5	−2	25	4	144	25
6	7	−1	0	1	0	36	49
合	計	合計	合計	40	10	285	255

$\overline{x}=\dfrac{1}{5}\boxed{35}=\boxed{7}$，　$\overline{y}=\dfrac{1}{5}\boxed{35}=\boxed{7}$　答

$s_x{}^2=\dfrac{1}{5}\boxed{40}=\boxed{8}$，　$s_y{}^2=\dfrac{1}{5}\cdot\boxed{10}=\boxed{2}$　答

$s_x=\sqrt{\boxed{8}}\fallingdotseq\boxed{2.8}$，　$s_y=\sqrt{\boxed{2}}\fallingdotseq\boxed{1.4}$　答

別解　$s_x{}^2=\dfrac{1}{5}\cdot\boxed{285}-\boxed{7}^2=\boxed{8}$，$s_y{}^2=\dfrac{1}{5}\cdot\boxed{255}-\boxed{7}^2=\boxed{2}$

(2)　x，y のデータについて，標準偏差によってデータの平均値からの散らばりの度合いを比較せよ。

(2)　(1) から　$s_x>s_y$

ゆえに，x のデータの方が，平均値からの散らばりの度合いが大きいと考えられる。

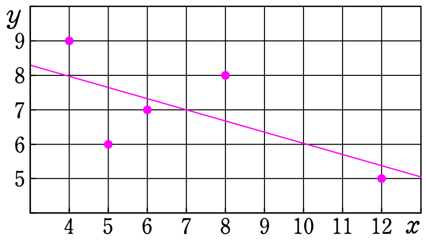
組　　番　名前

1 右の表は，ある製品を成型できる２台の工作機械 X，Y の１時間あたりのそれぞれの不良品の数 x，y を５時間にわたって調べたものである。(単位は個)
これをもとに，以下の各問いに答えよ。

x	5	4	8	12	6
y	6	9	8	5	7

(1)　散布図（相関図）をかけ。
横軸を x，縦軸を y として
右の図に● で表せ。

x	5	4	8	12	6
y	6	9	8	5	7



(2)　変量 x の平均値は
$\overline{x}$ ＝ $\boxed{7}$ である。

(3)　変量 y の平均値は
$\overline{y}$ ＝ $\boxed{7}$ である。

(4)　2つの変量の関係を調べるために以下のような表を作った。
空欄を補え。

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})(y-\overline{y})$	$(x-\overline{x})^2$	$(y-\overline{y})^2$
5	6	−2	−1	2	4	1
4	9	−3	2	−6	9	4
8	8	1	1	1	1	1
12	5	5	−2	−10	25	4
6	7	−1	0	0	1	0
合	計	合計	合計	−13	40	10

(5)　変量 x の分散は　s_x^2 ＝ $\boxed{8}$ であり，
標準偏差は　s_x ＝ $\boxed{2}\sqrt{\boxed{2}}$ ≒ $\boxed{2.8}$ である。
(小数点以下1位まで)

(6)　変量 y の分散は　s_y^2 ＝ $\boxed{2}$ であり，
標準偏差は　s_y ＝ $\sqrt{\boxed{2}}$ ≒ $\boxed{1.4}$ である。
(小数点以下1位まで)

x の分散
$s_x^2=\dfrac{1}{n}\sum_{k=1}^{n}(x_k-\overline{x})^2$
＝偏差の2乗の平均＝$\dfrac{10}{5}=2$

y の分散
$s_y^2=\dfrac{1}{n}\sum_{k=1}^{n}(y_k-\overline{y})^2$
＝偏差の2乗の平均＝$\dfrac{40}{5}=8$

別解　分散の求め方
$s_x{}^2=\dfrac{1}{5}(5^2+4^2+8^2+12^2+6^2)-7^2=8$，$s_y{}^2=\dfrac{1}{5}(6^2+9^2+8^2+5^2+7^2)-7^2=2$

(7)　x，y のデータについて，標準偏差によってデータの平均値からの散らばりの度合いを比較せよ。

(5)，(6) から　$s_x>s_y$　ゆえに，x のデータの方が，
平均値からの散らばりの度合いが大きいと考えられる。

> ２つの変量 x，y からなるデータとして n 個の値の組
> $(x_1,\ y_1)$，　$(x_2,\ y_2)$，　……，$(x_n,\ y_n)$
> **x の偏差と y の偏差の積の平均値**
> $\dfrac{1}{n}\{(x_1-\overline{x})(y_1-\overline{y})+(x_2-\overline{x})(y_2-\overline{y})+\cdots\cdots+(x_n-\overline{x})(y_n-\overline{y})\}$
> を x と y の **共分散** という。

(8)　変量 x と y の共分散は　s_{xy} ＝ − $\boxed{2}$. $\boxed{6}$ である。
(小数点以下1位まで)

共分散　$s_{xy}=\dfrac{1}{n}\{(x_1-\overline{x})(y_1-\overline{y})+(x_2-\overline{x})(y_2-\overline{y})+\cdots\cdots+(x_n-\overline{x})(y_n-\overline{y})\}$
＝x の偏差と y の偏差の積 $(x_k-\overline{x})(y_k-\overline{y})$ の平均値＝$\dfrac{-13}{5}=-2.6$

> ２つの変量 x，y からなるデータとして n 個の値の組
> $(x_1,\ y_1)$，　$(x_2,\ y_2)$，　……，$(x_n,\ y_n)$　があり
> 平均値をそれぞれ $\overline{x}$，$\overline{y}$，　標準偏差をそれぞれ s_x，s_y，
> 共分散を S_{xy} とするとき，
> $r=\dfrac{s_{xy}}{s_x s_y}$　を x と y の相関係数という。

(8)　変量 x と y の相関係数は
r ＝ − $\boxed{0}$. $\boxed{6}$ $\boxed{5}$ である。(小数点以下2位まで)

$r=\dfrac{s_{xy}}{s_x s_y}=\dfrac{-2.6}{\sqrt{8}\times\sqrt{2}}=-0.65$

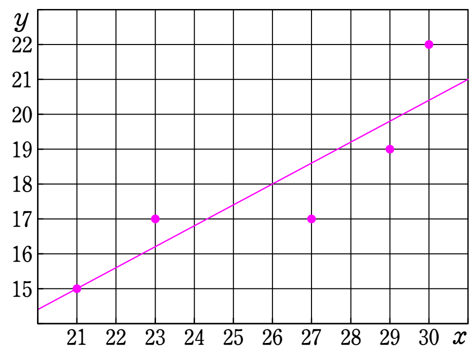
(9)　変量 x と y の相関関係について，正しいものを下記①～③のうちから一つ選ぶと $\boxed{②}$ である。
①　正の相関がある　②　負の相関がある　③　相関がない

2 2つの変量 x，y について，
相関関係を調べよ。

	①	②	③	④	⑤
x	21	27	29	23	30
y	15	17	19	17	22

(1)　散布図（相関図）をかけ。
横軸を x，縦軸を y として
右の図に ●で表せ。

x	21	27	29	23	30
y	15	17	19	17	22



(2)　x と y の相関の正負や
強弱を調べるために，
次のような表を作る。
ただし，$\overline{x}=\dfrac{1}{5}\times130=26$，
$\overline{y}=\dfrac{1}{5}\times90=18$ である。

	x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})(y-\overline{y})$	$(x-\overline{x})^2$	$(y-\overline{y})^2$
①	21	15	−5	−3	15	25	9
②	27	17	1	−1	−1	1	1
③	29	19	3	1	3	9	1
④	23	17	−3	−1	3	9	1
⑤	30	22	4	4	16	16	16
計	130	90	0	0	a	b	c

a ＝ $\boxed{36}$, b ＝ $\boxed{60}$, c ＝ $\boxed{28}$

(1)　変量 x の分散は s_x^2 ＝ $\boxed{12}$ ，標準偏差は s_x ＝ $\sqrt{\boxed{12}}$

(2)　変量 y の分散は s_y^2 ＝ $\boxed{5.6}$ ，標準偏差は s_y ＝ $\sqrt{\boxed{5.6}}$
(小数点以下1位まで)　($\sqrt{\ }$ の中は小数点以下1位まで)

(3)　変量 x と y の共分散は　s_{xy} ＝ $\boxed{7.2}$ (小数点以下1位まで)

(小数点以下1位まで)
(4)　相関係数は　$r=\dfrac{s_{xy}}{s_x s_y}=\dfrac{\boxed{7.2}}{\sqrt{\boxed{12}}\sqrt{\boxed{5.6}}}$ ≒0.878
($\sqrt{\ }$ の中は小数点以下1位まで)

> 別解　相関係数は上の表の値 a，b，c を用いた式
> $r=\dfrac{a}{\sqrt{bc}}$　で求めることもできる。

数学Ⅰ　データの分析　演習問題

新編数学A　P.170〜198

	組		番　名前	

1 次のデータは，ある店における商品 A，B の 10日間の販売数である。

商品 A　　12，9，5，7，13，12，10，6，7，8

商品 B　　10，15，16，20，8，12，13，22，16，11

(1)　階級の幅を 2 個として，**度数分布表**を作れ。ただし，階級は 4 個から区切り始めるものとする。

A	12	9	5	7	13	12	10	6	7	8
B	10	15	16	20	8	12	13	22	16	11

階級	度数 A	度数 B
4以上 6未満	1	0
6 〜 8	3	0
8 〜 10	2	1
10 〜 12	1	2
12 〜 14	3	2
14 〜 16	0	1
16 〜 18	0	2
18 〜 20	0	0
20 〜 22	0	1
22 〜 24	0	1
計	10	10

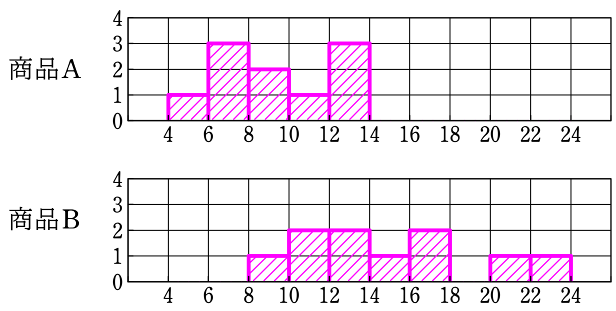
(2)　12以上14未満の階級について，

階級値は $\boxed{13}$ であり，

商品Aの度数は $\boxed{3}$，

商品Bの度数は $\boxed{2}$ である。

(3)　(1) で作った度数分布表をもとにして，**ヒストグラム**をかけ。



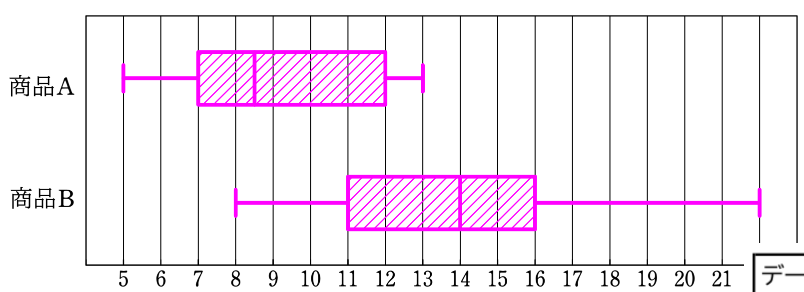
(4)　データを小さい順に並べ替えよ。

A	5	6	7	7	8	9	10	12	12	13
B	8	10	11	12	13	15	16	16	20	22

(5)　**(4) で並べ替えたデータをもとに，商品A，商品Bのそれぞれについて，次の各値を求め空欄を補え。**（小数点以下1位まで）

商品	A	B
平均値	8.9	14.3
中央値	8.5	14.0
最頻値	7.0 , 12.0	16.0
最大値	13.0	22.0
最小値	5.0	8.0
第1四分位数	7.0	11.0
第2四分位数	8.5	14.0
第3四分位数	12.0	16.0
四分位範囲	5.0	5.0
四分位偏差	2.5	2.5

(6)　(5) で作った数値をもとに，これらのデータの箱ひげ図を並べてかけ。



2 右の表は，5 人の生徒に 10 点満点の数学と英語の小テストを行った結果である。数学を x，英語を y で表すものとする。

生徒	A	B	C	D	E
数学 x	7	5	9	6	3
英語 y	6	7	8	10	4

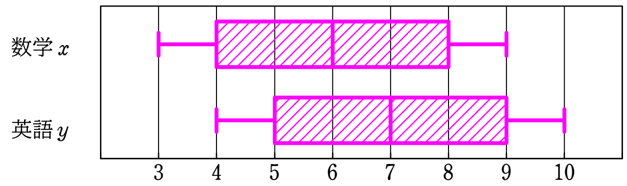
(1)　データを小さい順に並べ替えよ。

数学 x	3	5	6	7	9
英語 y	4	6	7	8	10

(2)　**(1) で並べ替えたデータをもとに，数学 x，英語 y のそれぞれについて，次の各値を求め空欄を補え。**

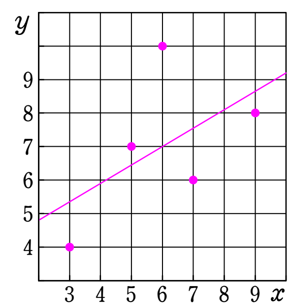
	数学 x	英語 y
平均値	6	7
中央値	6	7
最大値	9	10
最小値	3	4
第1四分位数	4	5
第2四分位数	6	7
第3四分位数	8	9
四分位範囲	4	4
四分位偏差	2	2

(3)　(2) の数値をもとに，これらのデータの箱ひげ図を並べてかけ。



(4)　散布図（相関図）をかけ。
数学を x 軸，英語を y 軸で表すものとする。（右の図に●で表せ。）

共分散 2.20000
相関係数 0.55000



(5)　2つの変量の関係を調べるために以下のような表を作った。空欄を補え。

	x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})(y-\overline{y})$	$(x-\overline{x})^2$	$(y-\overline{y})^2$
A	7	6	1	−1	−1	1	1
B	5	7	−1	0	0	1	0
C	9	8	3	1	3	9	1
D	6	10	0	3	0	0	9
E	3	4	−3	−3	9	9	9
計	30	35	0	0	a	b	c

$a=\boxed{11}$，$b=\boxed{20}$，$c=\boxed{20}$

(6)　変量 x の分散は $s_x^2=\boxed{4}$，標準偏差は $s_x=\boxed{2}$

(7)　変量 y の分散は $s_y^2=\boxed{4}$，標準偏差は $s_y=\boxed{2}$

(8)　変量 x と y の共分散は　$s_{xy}=\boxed{2.2}$（小数点以下1位まで）

（小数点以下1位まで）

(9)　相関係数は　$r=\dfrac{s_{xy}}{s_x s_y}=\dfrac{\boxed{2.2}}{\boxed{2}\times\boxed{2}}=\boxed{0.55}$

（小数点以下2位まで）

別解　相関係数は上の表の値 a，b，c を用いた式

$r=\dfrac{a}{\sqrt{bc}}$　で求めることもできる。

(10)　変量 x と y の相関関係について，正しいものを下記①〜③のうちから一つ選ぶと $\boxed{①}$ である。

①　正の相関がある　　②　負の相関がある　　③　相関がない

組　　番　名前

1 次の10個のデータ 3，2，7，3，6，3，5，3，8，5 において，各問いに答えよ。（解答の様式は下記点線枠内を参考にせよ。）

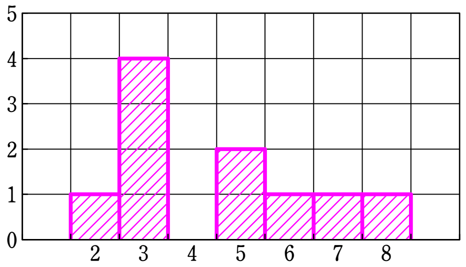
(1) 小さい順に並べよ。

データ	2	3	3	3	3	5	5	6	7	8

(2) 度数分布表の空欄を補え。

階級	度数
2	1
3	4
4	0
5	2
6	1
7	1
8	1
計	10

(3) ヒストグラムをかけ。



(4) 平均値は $\boxed{4}$. $\boxed{5}$ である。（小数点以下1位まで）

(5) 中央値は $\boxed{4}$ である。

(6) 最頻値は $\boxed{3}$ である。

(7) 第１四分位数は $\boxed{3}$ である。

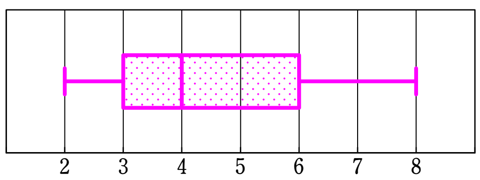
(8) 第２四分位数は $\boxed{4}$ である。

(9) 第３四分位数は $\boxed{6}$ である。

(10) 四分位範囲は $\boxed{3}$ である。

(11) 四分位偏差は $\boxed{1}$. $\boxed{5}$ である。（小数点以下1位まで）

(12) 箱ひげ図をかけ。（ただし，平均値は記入しなくてよい。）



点線枠内は解答様式の見本である。

データ	1	2	3	3	5

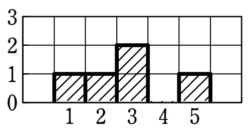
平均値は $\boxed{2}$. $\boxed{8}$

（小数点以下1位まで）

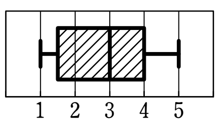
度数分布表

階級	度数
1	1
2	1
3	2
4	0
5	1
計	5

ヒストグラム



箱ひげ図



2 次の2つの変量 x と y についてのデータがある。

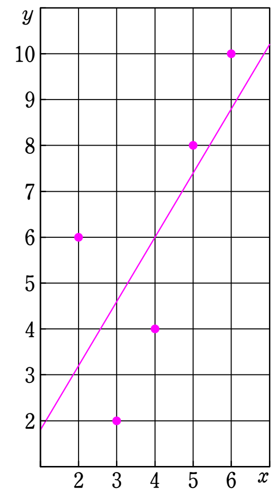
x	3	4	6	2	5
y	2	4	10	6	8

x	3	4	6	2	5
y	2	4	10	6	8

これをもとに，以下の各問いに答えよ。

(1) 相関図（散布図）をかけ。

（右の図に●で表せ。）



(2) 変量 x の平均値は

$\overline{x} = \boxed{4}$ である。

(3) 変量 y の平均値は

$\overline{y} = \boxed{6}$ である。

(4) 2つの変量の関係を調べるために以下のような表を作った。空欄を補え。

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})(y-\overline{y})$	$(x-\overline{x})^2$	$(y-\overline{y})^2$
3	2	−1	−4	4	1	16
4	4	0	−2	0	0	4
6	10	2	4	8	4	16
2	6	−2	0	0	4	0
5	8	1	2	2	1	4
合	計	合計	合計	14	10	40

(5) 変量 x の分散は $s_x^2 = \boxed{2}$ である。

x の分散

$s_x^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k-\overline{x})^2$

$=$ 偏差の2乗の平均 $= \frac{10}{5} = 2$

(6) 変量 y の分散は $s_y^2 = \boxed{8}$ である。

y の分散

$s_y^2 = \frac{1}{n}\sum_{k=1}^{n}(y_k-\overline{y})^2$

$=$ 偏差の2乗の平均 $= \frac{40}{5} = 8$

(7) 変量 x と y の共分散は $s_{xy} = \boxed{2}$. $\boxed{8}$ である。

（小数点以下1位まで）

共分散 $s_{xy} = \frac{1}{n}\{(x_1-\overline{x})(y_1-\overline{y}) + (x_2-\overline{x})(y_2-\overline{y}) + \cdots\cdots + (x_n-\overline{x})(y_n-\overline{y})\}$

$=$ x の偏差と y の偏差の積 $(x_k-\overline{x})(y_k-\overline{y})$ の平均値 $= \frac{14}{5} = 2.8$

(8) 変量 x と y の相関係数は

$r = \boxed{0}$. $\boxed{7}$ である。（小数点以下1位まで）

$r = \frac{s_{xy}}{s_x s_y} = \frac{2.8}{\sqrt{2}\times\sqrt{8}} = 0.7$

$= \frac{x\text{ の偏差と }y\text{ の偏差の積の和}}{\sqrt{x\text{ の偏差の2乗の和}}\times\sqrt{y\text{ の偏差の2乗の和}}} = \frac{14}{\sqrt{10}\times\sqrt{40}} = \frac{14}{20} = 0.7$

(9) 変量 x と y の相関関係について，正しいものを下記①〜③のうちから一つ選ぶと $\boxed{①}$ である。

① 正の相関がある

② 負の相関がある

③ 相関がない

1 仮説検定の考え方（教科書を一部変更）

旧製品 A と新製品Bの評価を調査した。
無作為に選んだ 30 人　に 2 つの製品 A，B の評価を聞いたところ，
「70 % にあたる 21 人が B がよい」　と回答した。
この回答から，「B の方がよい」　と判断してよいか。
基準となる 確率を 0.05 として検定せよ。
（この基準確率を有意水準と呼び，「これ以下の確率は起こらない」と考えます。）

解説

step1.　もとの主張を「B の方がよい。」とする。

step2.　反対の仮定を立てる（帰無仮説）

「B の方がよいとは言えない。」（偶然だ）

解説 注意　つまり，もとの仮定を正しいと言うために
（反対の仮定を捨てるために），
反対の仮定が起こる確率をデータに基づいて算出し
検定結果の確率＜基準の確率　0.05
「検定結果が基準より小さいことを示して
そんなことは起こらないぢゃないか！」
として反対の仮定を捨てます。（帰無仮説の棄却）

step3.　与えられたデータに基づいて，調査結果を検証する。

コイン投げ30回を200セット繰り返したデータ

表の回数	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	計
度数	2	3	3	12	16	22	22	31	31	22	14	14	4	2	1	1	200

20回以上表が出たのは200セットのうち $2+1+1=4$ セットであり、
相対度数は $\dfrac{4}{200}=0.02$ ＜　基準確率 0.05　である。

step4.　結論を書く。

そもそもstep.2 の仮定が正しくなかったと考えて
仮説を捨てます。（帰無仮説の棄却）
答　よって，もとの主張は正しい，
つまり B の方がよいと判断できる。

2 仮説検定［714新編　数学Ⅰ　例12］

同じ調査で　19 人が B がよい　と回答したとしよう。
B の方がよい　とする主張を，基準となる確率を 0.05
として検定せよ。ただし，以下のデータをもとにすること。

表の回数	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	計
度数	2	3	3	12	16	22	22	31	31	22	14	14	4	2	1	1	200

解答[1] の主張に反する仮定を次のように立てる。
[2]　A，B のどちらの回答も全くの偶然で起こる
[2] の仮定は，公正なコインを 30 回投げて表の出た回数を記録する実験にあてはめることができる。ここでは，コインの表が出る場合を，B と回答する場合とする。このコイン投げの実験結果を利用すると，19 回以上表が出る場合の相対度数は

$$\frac{14+4+2+1+1}{200}=\frac{22}{200}=0.11$$

検定結果 0.11 ＞ 基準 0.05
ありうることだ
＝仮説を棄却できない

これは　0.05 より大きいから，主張 [2] は否定できない。
よって，　答　B の方が書きやすいと評価されるとは判断できない。

3 仮説検定の考え方：さいころの1の目が出にくいかの判断［3TRIAL数学Ⅰ　問題296］

あるさいころを 30 回投げたところ 1 の目が 1 回しか出なかった。
このさいころは 1 の目が出にくいと判断してよいか。仮説検定の考え方を用い，基準となる確率を 0.05 として考察せよ。
ただし，公正なさいころを 30 回投げて 1 の目が出た回数を記録する実験を 300 セット行ったところ次の表のようになったとし，この結果を用いよ。

1 の目が出た回数	0	1	2	3	4	5	6	7	8	9	10	11	計
度数	2	7	20	44	58	56	45	40	17	9	1	1	300

解答　[1]　1 の目が出にくい　と判断できるかを考察するために，
[1] に反する次の仮定を立てる。
[2]　どの目も等しい確率で出る
さいころ投げの実験結果を利用すると，1 の目が 0 回または 1 回しか出ない場合の相対度数は

$$\frac{2+7}{300}=\frac{9}{300}=0.03$$

検定結果 0.03 ＜ 基準 0.05
等しいとする仮定には無理がある。＝仮説を棄却する。
元の主張が正しいと判断。

これは 0.05 より小さいから，[2] の仮定は正しくなかったと考えられる。
したがって，
答　主張 [1] は正しい，すなわち 1 の目は出にくいと判断できる。

4 仮説検定による判断：コイン投げの実験結果利用［白チャート数学Ⅰ　例題155］

ある会社では，既に販売しているボールペン A を改良したボールペン B を開発した。書きやすさを評価してもらうために，無作為に選んだ 20 人に，A と B のどちらが書きやすいかのアンケートを行った結果，15 人が B と回答した。
このアンケート結果から，B の方が書きやすいと消費者から評価されていると判断してよいか。
基準となる確率を 0.05 とし，次のコイン投げの実験の結果を利用して考察せよ。

実験　公正な 1 枚のコインを投げる。そして，コイン投げを 20 回行うことを 1 セットとし，1 セットで表の出た枚数を記録する。
この実験を 200 セット繰り返したところ，次の表のような結果となった。

表の枚数	5	6	7	8	9	10	11	12	13	14	15	16	計
度数	4	10	15	19	27	33	29	26	21	12	3	1	200

解答　「主張：B の方が書きやすい」に対する次の仮説を立てる。
仮説：A と B の書きやすさに差はない。
つまり，A と回答する場合と B と回答する場合が
半々の確率で起こる。
ここで，コインの表が出る場合を，B と回答する場合とする。
コイン投げの実験結果において，15 枚以上表が出たのは，200 セットのうち，$3+1=4$
セットであり，相対度数は
$4\div200=0.02$
これは基準の確率 0.05 より小さい。
したがって，仮説は正しくないと考えられる。
よって，最初の主張は正しい。
つまり，「B の方が書きやすい」と評価されていると判断してよい。

検定結果 0.02 ＜ 基準 0.05
差がないとする仮定には無理がある。＝仮説を棄却する。
元の主張が正しいと判断。

B の方が書きやすい」と評価されていると判断してよい

課題　偏差値ってなんなの？　　　　　月　　日

生徒：わー、この前のテスト、難しくて 65 点だった。

先生：同じ 100 点満点のテストでも、難しければみんな点数が悪いし、やさしければいい点を取る。で、平均点は？

生徒：47 点。

先生：なーんだ。だったら悪くない。ところで、問題は難しくても簡単でも、大体何番目ぐらいの成績が分かる「偏差値」を知っとるか？

生徒：よく聞くが、意味は分からん。

先生：大勢が受けたテストの結果を点数ごとの人数でグラフにすると、普通は「正規分布」というなめらかな山の形になると考えられている。でも、テストが簡単だと山の頂上が点の高いほうに、難しければ低いほうに、かたよるし、平均点が同じでも山がとんがっていたり、緩やかだったりすることもある。

生徒：つまり点数だけでは、実力がよく分からないんだ。

先生：そこで「偏差値」が考えられた。どんなテストの結果でもグラフが標準的な同じ山の形になるように、点数を調整するんだ。

生徒：どう調整するの。

先生：まず平均点の偏差値を５０にする。そして「標準偏差」というものさしで、自分の点数が平均点より標準偏差のいくつ分高いか、低いかを計算する。標準偏差が一つ分ずれるごとに１０ずつ、50 に加えたり減らしたりするのだ。

生徒：標準偏差って？

先生：簡単にいえば、点数のグラフの山のすそのがどのくらい遠くまで広がっているかを示す値だ。テストの結果が 0 点から 100 点まで広く分布していると標準偏差は大きいし、平均点のまわりに集中していると小さくなる。

生徒：ぼくの偏差値も計算できるの。

先生：仮にテストの標準偏差が 15 点だとすると、生徒の 65 点は平均点より 18 点高い。これは、18÷15 で標準偏差の 1.2 個分だ。標準偏差 1 個で偏差値は 10 ずれるから、1.2 個分だと１２ずれる。だから 50＋12 で、生徒の偏差値は６２！

生徒：ふーん、それで成績は何番目くらい？

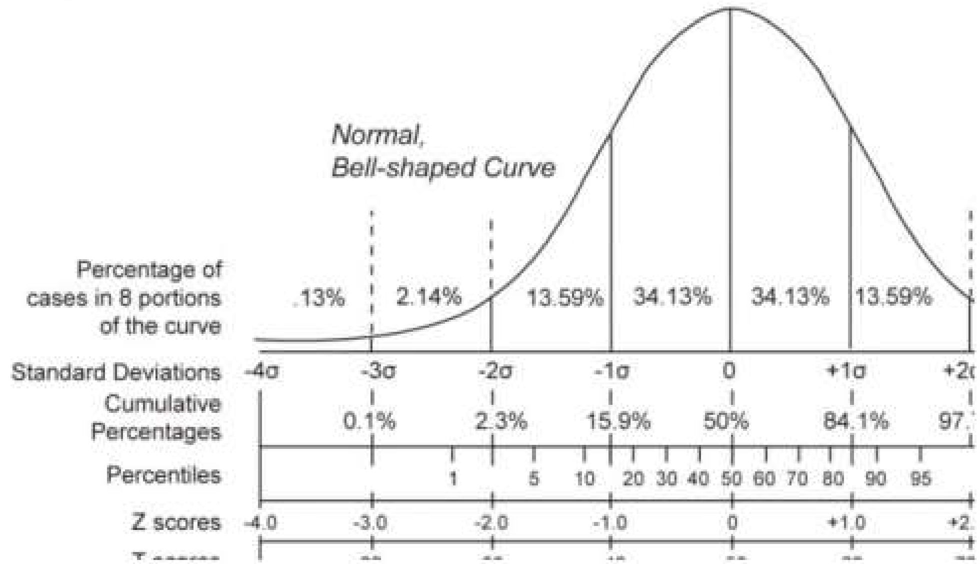
先生：成績全体が正規分布しているとすると、偏差値 60 以上は上位約 16%に入る計算になる。偏差値 62 なら 100 人のうち十何番目かのあたりだ。結構自信を持ってよい。

生徒：やったー。

先生：でも、「偏差値は、学力による人間の序列化だ」という批判もあるし、このごろのテストは結果が二つの山に分かれる形になることが多いから、正規分布を前提にする偏差値では成績が正しく表せないという意見もある。たしかに偏差値の気にしすぎはよくない。一つの目安ぐらいの気持ちで、偏差値の仕組みを理解して参考ににすればよいのかも。

生徒：なるほど。　　　　　【朝日新聞　DO 科学　2009.9.19.より】

資料



問題

次の表は、10 人の生徒が 100 点満点のテストを受けた結果の表です。この表から各生徒の偏差値を求めよ。

生徒	A	B	C	D	E	F	G	H	I	J	計
得点	50	90	60	60	40	100	40	40	50	70	
偏差											—
偏差の２乗											
偏差÷標準偏差 ×１０											—
偏差値											—

＊平均点：（得点の計）÷人数

＊分散：（偏差の２乗の計）÷人数

＊標準偏差：分散の正の平方根

組　　番　名前

偏差値ってなんなの？　解答

生徒	A	B	C	D	E	F	G	H	I	J	計
得点	50	90	60	60	40	100	40	40	50	70	600
偏差	-10	30	0	0	-20	40	-20	-20	-10	10	0
偏差の2乗	100	900	0	0	400	1600	400	400	100	100	4000
偏差÷標準偏差×10	-5	15	0	0	-10	20	-10	-10	-5	5	0
偏差値	45	65	50	50	40	70	40	40	45	55	

平均	60
s^2 （分散）	400
s=√s^2 （標準偏差）	20

偏差値はZ値とも言われるが、次の式で計算する。

$$Z=\frac{10(x-m)}{\sigma}+50$$

（xは個々の得点、mは平均、σは標準偏差）

データの分析 8 数学Ⅰ（714）

データの分析 9 数学Ⅰ（714）