データの分析

1 次のデータの平均値，最頻値，中央値を求めよ。

(1) 1，2，4，4，5，6，7，7，7，8

(2) 4，13，6，9，3，9，6，13，9，6，3，14，9

解答 (1) 平均値 5.1，最頻値 7，中央値 5.5

(2) 平均値 8，最頻値 9，中央値 9

解説

(1) データの大きさは 10 であるから，平均値は

$$\frac{1}{10}(1+2+4+4+5+6+7+7+7+8)=\frac{1}{10}\times 51=5.1$$

最頻値は 7

中央値は $\frac{5+6}{2}=5.5$

(2) データを小さい順に並べると

3 3 4 6 6 6 9 9 9 9 13 13 14

データの大きさは 13 であるから，平均値は

$$\frac{1}{13}(3+3+4+6+6+6+9+9+9+9+13+13+14)=\frac{1}{13}\times 104=8$$

最頻値は 9

中央値は 9

2 次のデータの第 1 四分位数，第 2 四分位数，第 3 四分位数を求めよ。

(1) 12，35，47，59，68，73，74，79，87，97

(2) 2，7，10，14，22，33，48，55，71，84，91，96，98

解答 第 1 四分位数，第 2 四分位数，第 3 四分位数の順に

(1) 47，70.5，79 (2) 12，48，87.5

解説

第 1 四分位数 Q_1，第 2 四分位数 Q_2，第 3 四分位数 Q_3 は，それぞれ次のようになる。

(1) $Q_2=\frac{68+73}{2}=70.5$，$Q_1=47$，$Q_3=79$

(2) $Q_2=48$，$Q_1=\frac{10+14}{2}=12$，$Q_3=\frac{84+91}{2}=87.5$

3 次のデータ A，B のそれぞれについて，四分位範囲と四分位偏差を求めよ。また，データの散らばりの度合いが大きいのは A，B のどちらと考えられるか。得られた四分位範囲によって比較せよ。

A 2，13，14，13，11，18，13，5，10，14，5，7

B 5，9，15，13，7，14，12，7，10，13，4，8

解答 四分位範囲，四分位偏差の順に A：7.5，3.75；B：6，3

データの散らばりの度合いが大きいのは A

解説

A のデータを小さい方から順に並べると

2 5 5 7 10 11 13 13 13 14 14 18

A のデータについて，第 1 四分位数，第 3 四分位数はそれぞれ

$Q_1=\frac{5+7}{2}=6$，$Q_3=\frac{13+14}{2}=13.5$

よって，四分位範囲は $Q_3-Q_1=13.5-6=7.5$

四分位偏差は $\frac{Q_3-Q_1}{2}=\frac{7.5}{2}=3.75$

B のデータを小さい方から順に並べると

4 5 7 7 8 9 10 12 13 13 14 15

B のデータについて，第 1 四分位数，第 3 四分位数はそれぞれ

$Q_1=\frac{7+7}{2}=7$，$Q_3=\frac{13+13}{2}=13$

よって，四分位範囲は $Q_3-Q_1=13-7=6$

四分位偏差は $\frac{Q_3-Q_1}{2}=\frac{6}{2}=3$

A の方が四分位範囲が大きいから，データの散らばりの度合いが大きいと考えられるのは，A である。

4 次のデータは，6 人の生徒のハンドボール投げの記録 x である。

26，25，32，28，32，25 (m)

(1) 偏差の 2 乗の平均値を求めることにより，分散 s^2 を求めよ。

(2) 標準偏差 s を求めよ。

解答 (1) 9 (2) 3 m

解説

(1) 平均値 $\overline{x}$ は

$$\overline{x}=\frac{1}{6}(26+25+32+28+32+25)=\frac{1}{6}\times 168=28 \text{ (m)}$$

偏差の 2 乗は右の表のようになる。

x	26	25	32	28	32	25	計 168
$(x-\overline{x})^2$	4	9	16	0	16	9	計 54

よって，分散 s^2 は

$$s^2=\frac{1}{6}\times 54=9$$

(2) 標準偏差 s は $s=\sqrt{s^2}=\sqrt{9}=3$ (m)

5 5 個の値 50，70，90，80，50 からなる変量 x のデータがある。

(1) 各値の 2 乗の平均値を求めることにより，分散 s^2 を求めよ。

(2) 標準偏差 s を求めよ。

解答 (1) 256 (2) 16

解説

(1) 変量 x，x^2 のデータの値は右の表のようになる。

x	50	70	90	80	50	計 340
x^2	2500	4900	8100	6400	2500	計 24400

$\overline{x}=\frac{1}{5}\times 340=68$，

$\overline{x^2}=\frac{1}{5}\times 24400=4880$ である。

よって，分散 s^2 は $s^2=\overline{x^2}-(\overline{x})^2=4880-68^2=256$

(2) 標準偏差 s は $s=\sqrt{s^2}=\sqrt{256}=16$

6 次のデータの平均値，最頻値，中央値を求めよ。

6，7，7，7，9，10，11，11，14，16

解答 平均値 9.8，最頻値 7，中央値 9.5

解説

データの大きさは 10 であるから，平均値は

$$\frac{1}{10}(6+7+7+7+9+10+11+11+14+16)=\frac{1}{10}\times 98=9.8$$

最頻値は 7

中央値は $\frac{9+10}{2}=9.5$

7 次の変量 x のデータについて，次の問いに答えよ。

5，1，8，10，4，8

(1) 偏差の 2 乗の平均値を求めることにより，分散 s^2 を求めよ。

(2) 各値の 2 乗の平均値 $\overline{x^2}$ を求めることにより，分散 s^2 を求めよ。

(3) 標準偏差 s を求めよ。

解答 (1) 9 (2) 9 (3) 3

解説

(1) 平均値 $\overline{x}$ は $\overline{x}=\frac{1}{6}(5+1+8+10+4+8)=\frac{1}{6}\times 36=6$

偏差の 2 乗は右の表のようになるから，

x	5	1	8	10	4	8	計 36
$(x-\overline{x})^2$	1	25	4	16	4	4	計 54

分散 s^2 は $s^2=\frac{1}{6}\times 54=9$

(2) $\overline{x^2}=\frac{1}{6}(5^2+1^2+8^2+10^2+4^2+8^2)=\frac{1}{6}\times 270=45$

分散 s^2 は $s^2=\overline{x^2}-(\overline{x})^2=45-6^2=9$

(3) 標準偏差 s は $s=\sqrt{s^2}=\sqrt{9}=3$

8 (1) 次のデータは，生徒 10 人のテストの得点である。このデータの平均値を求めよ。

69	73	33	41	78	51	5	57	36	75

(2) 次の表は，ある店舗での洋服のサイズ別の 1 週間の販売数である。このデータの最頻値を求めよ。

サイズ(cm)	90	100	110	120	130	140	150	160	計
販売数	29	30	28	22	41	26	37	12	225

(3) 次の 5 個のデータの中央値を求めよ。

46	67	14	88	18

(4) 次の 6 個のデータの中央値を求めよ。

68	50	90	78	57	7

解答 (1) 51.8 点 (2) 130 cm (3) 46 (4) 62.5

解説

(1) 得点のデータの平均値 $\overline{x}$ は

$$\overline{x}=\frac{1}{10}(69+73+33+41+78+51+5+57+36+75)$$

$$=\frac{1}{10}\times 518=51.8 \text{ (点)}$$

(2) 洋服の販売数の最大値は 41 である。

よって，このデータの最頻値は 130 cm である。

(3) このデータを小さい方から並べると

14 18 46 67 88

よって，中央値は 46

(4) このデータを小さい方から並べると

7 50 57 68 78 90

よって，中央値は　$\dfrac{57+68}{2}=62.5$

9　次のデータの平均値 $\overline{x}$ を求めよ。

(1)　16，34，36，33，28

(2)　62，80，58，20，57，37，15

(3)　28，1，17，22，4，29，0，37

解答　(1)　29.4　　(2)　47　　(3)　17.25

解説

(1)　$\overline{x}=\dfrac{1}{5}(16+34+36+33+28)=\dfrac{147}{5}=29.4$

(2)　$\overline{x}=\dfrac{1}{7}(62+80+58+20+57+37+15)=\dfrac{329}{7}=47$

(3)　$\overline{x}=\dfrac{1}{8}(28+1+17+22+4+29+0+37)=\dfrac{138}{8}=17.25$

10　次のデータの中央値を求めよ。

(1)　22，14，97，48，8

(2)　81，87，77，79，39，86，35

(3)　33，40，64，59，60，62，20，91

(4)　70，95，22，86，0，91，88，2，15，91

解答　(1)　22　　(2)　79　　(3)　59.5　　(4)　78

解説

(1)　データを小さい方から並べると

8，14，22，48，97

よって，中央値は　22

(2)　データを小さい方から並べると

35，39，77，79，81，86，87

よって，中央値は　79

(3)　データを小さい方から並べると

20，33，40，59，60，62，64，91

よって，中央値は　$\dfrac{59+60}{2}=59.5$

(4)　データを小さい方から並べると

0，2，15，22，70，86，88，91，91，95

よって，中央値は　$\dfrac{70+86}{2}=78$

11　次のデータは，生徒 8 人の教科 A，B のテストの得点である。A，B の得点の平均値をそれぞれ求めよ。

A　67，72，91，53，39，74，83，77　(点)

B　61，59，78，58，65，70，57，80　(点)

解答　教科 A は 69.5 点，教科 B は 66 点

解説

教科 A のデータの平均値は

$\dfrac{1}{8}(67+72+91+53+39+74+83+77)=\dfrac{556}{8}=69.5$ (点)

教科 B のデータの平均値は

$\dfrac{1}{8}(61+59+78+58+65+70+57+80)=\dfrac{528}{8}=66$ (点)

12　次のデータの中央値を求めよ。

(1)　51，39，20，11，38，50，20

(2)　74，66，29，50，33，76，8，91，96，98

解答　(1)　38　　(2)　70

解説

(1)　データを小さい方から並べると

11，20，20，38，39，50，51

よって，中央値は　38

(2)　データを小さい方から並べると

8，29，33，50，66，74，76，91，96，98

よって，中央値は　$\dfrac{66+74}{2}=70$

13　(1)　次のデータは，A 組と B 組における図書室での本の貸し出し冊数を月ごとに調べたものである。それぞれのデータの範囲を求めよ。また，データの散らばり具合が大きいのは A 組，B 組のどちらか。

A 組　27，19，15，32，9，23，38，41，17，21，31，29　(冊)

B 組　18，23，16，33，21，15，12，28，35，37，22，19　(冊)

(2)　次のデータは，6 人の生徒のハンドボール投げの記録 x である。

26，25，32，28，32，25　(m)

偏差の 2 乗の平均値を求めて，x の分散 s^2 と標準偏差 s を求めよ。

解答　(1)　A 組は 32 冊，B 組は 25 冊，散らばり具合が大きいのは A 組

(2)　$s^2=9$，$s=3$ (m)

解説

(1)　A 組のデータの範囲は　　$41-9=32$　(冊)

B 組のデータの範囲は　　$37-12=25$　(冊)

よって，データの散らばり具合が大きいのは A 組。

(2)　平均値 $\overline{x}$ は　$\overline{x}=\dfrac{1}{6}(26+25+32+28+32+25)=\dfrac{1}{6}\times168=28$

x	26	25	32	28	32	25	計168
$x-\overline{x}$	-2	-3	4	0	4	-3	
$(x-\overline{x})^2$	4	9	16	0	16	9	計54

表から分散 s^2 は　　$s^2=\dfrac{1}{6}\times54=9$

標準偏差 s は　$s=\sqrt{9}=3$ (m)

14　次のデータは，ある年の A さんと B さんが月ごとに購入した本の冊数を調べたものである。それぞれのデータの範囲を求めよ。また，データの散らばり具合が大きいのは A さん，B さんのどちらか。

A さん　3，4，6，8，5，4，2，9，4，5，3，3 (冊)

B さん　2，3，1，2，4，3，5，6，2，2，3，2 (冊)

解答　A さんは 7 冊，B さんは 5 冊，散らばり具合が大きいのは A さん

解説

A さんのデータの範囲は　$9-2=7$ (冊)

B さんのデータの範囲は　$6-1=5$ (冊)

よって，A さんの方がデータの散らばり具合が大きい。

15　次のデータは，10 人の数学の評価 x である。

3，7，9，6，3，5，6，2，4，5　(10 段階評価)

(1)　偏差の 2 乗の平均値を求めることにより，分散 s^2 を求めよ。

(2)　標準偏差 s を求めよ。

解答　(1)　4　　(2)　2

解説

(1)　平均値 $\overline{x}$ は

$\overline{x}=\dfrac{1}{10}(3+7+9+6+3+5+6+2+4+5)=\dfrac{50}{10}=5$

分散 s^2 は

$$s^2=\frac{1}{10}\{(3-5)^2+(7-5)^2+(9-5)^2+(6-5)^2+(3-5)^2+(5-5)^2+(6-5)^2+(2-5)^2+(4-5)^2+(5-5)^2\}$$

$$=\frac{1}{10}(4+4+16+1+4+0+1+9+1+0)=\frac{40}{10}=4$$

(2)　標準偏差 s は　$s=\sqrt{4}=2$

16　次のデータは，8 人の小テストの得点 x である。

7，5，7，6，8，7，10，6　(点)

x^2 の平均値 $\overline{x^2}$ を求めることにより，分散 s^2，標準偏差 s を求めよ。ただし，標準偏差は小数第 2 位を四捨五入せよ。

解答　$s^2=2$，$s=1.4$ (点)

解説

平均値 $\overline{x}$ は

$\overline{x}=\dfrac{1}{8}(7+5+7+6+8+7+10+6)=\dfrac{56}{8}=7$ (点)

x^2 の平均値 $\overline{x^2}$ は

$\overline{x^2}=\dfrac{1}{8}(7^2+5^2+7^2+6^2+8^2+7^2+10^2+6^2)=\dfrac{408}{8}=51$

分散 s^2 は　　$s^2=\overline{x^2}-(\overline{x})^2=51-7^2=2$

標準偏差 s は　$s=\sqrt{2}=1.41\cdots\cdots\fallingdotseq1.4$ (点)

17　(1)　次のデータの第 1 四分位数 Q_1，第 2 四分位数 Q_2，第 3 四分位数 Q_3 を求めよ。

5，7，12，19，21，30，33，36，40

(2)　次の変量 x，y のデータの四分位範囲と四分位偏差を求めよ。また，x，y のデータの散らばり具合の大きい方をいえ。

x	0	0	3	3	4	5	7	8	8	9
y	2	3	5	5	6	7	7	8	8	9

(3)　次のデータは，ある書店における書籍 A の 1 年間の販売数である。

12，14，11，20，9，17，10，12，17，21，15，16　(冊)

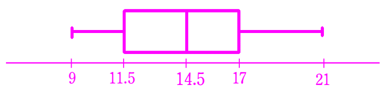
このデータの箱ひげ図をかけ。

解答　(1)　$Q_1=9.5$，$Q_2=21$，$Q_3=34.5$

(2)　x の四分位範囲 5，四分位偏差 2.5

y の四分位範囲 3，四分位偏差 1.5
散らばり具合が大きいのは x

(3)

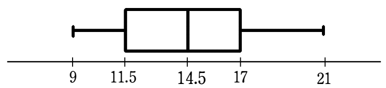


解説

(1)　この 9 個のデータの中央値は 21
すなわち，第 2 四分位数　$Q_2=21$
下半分 4 個のデータの中央値は　$\dfrac{7+12}{2}=9.5$
よって，第 1 四分位数　$Q_1=9.5$
上半分 4 個のデータの中央値は　$\dfrac{33+36}{2}=34.5$
よって，第 3 四分位数　$Q_3=34.5$

(2)　変量 x について，第 1 四分位数 $Q_1=3$，第 3 四分位数 $Q_3=8$
四分位範囲は　$Q_3-Q_1=8-3=5$
四分位偏差は　$\dfrac{Q_3-Q_1}{2}=\dfrac{5}{2}=2.5$
変量 y について，第 1 四分位数 $Q_1=5$，第 3 四分位数 $Q_3=8$
四分位範囲は　$Q_3-Q_1=8-5=3$
四分位偏差は　$\dfrac{Q_3-Q_1}{2}=\dfrac{3}{2}=1.5$
よって，変量 x の方が散らばり具合が大きい。

(3)　このデータを小さい方から並べると
9，10，11，12，12，14，15，16，17，17，20，21
このデータの最小値は 9，第 1 四分位数は $Q_1=\dfrac{11+12}{2}=11.5$，
中央値は $\dfrac{14+15}{2}=14.5$，第 3 四分位数は $Q_3=\dfrac{17+17}{2}=17$，
最大値は 21
箱ひげ図は，次のようになる。



18 次のデータの第 1 四分位数 Q_1，第 2 四分位数 Q_2，第 3 四分位数 Q_3 を求めよ。
(1)　9，15，17，22，29，31，40，43
(2)　8，11，20，23，24，31，42，44，55，58
(3)　10，12，18，21，31，33，42，45，48，52，56
(4)　3，9，12，14，34，38，41，42，45，52，54，59，64

解答　(1)　$Q_1=16$，$Q_2=25.5$，$Q_3=35.5$
(2)　$Q_1=20$，$Q_2=27.5$，$Q_3=44$
(3)　$Q_1=18$，$Q_2=33$，$Q_3=48$
(4)　$Q_1=13$，$Q_2=41$，$Q_3=53$

解説

(1)　データは 8 個であるから，中央値すなわち，
第 2 四分位数は　$Q_2=\dfrac{22+29}{2}=25.5$
下半分 4 個のデータの中央値すなわち，
第 1 四分位数は　$Q_1=\dfrac{15+17}{2}=16$
上半分 4 個のデータの中央値すなわち，
第 3 四分位数は　$Q_3=\dfrac{31+40}{2}=35.5$

(2)　データは 10 個であるから，中央値すなわち，
第 2 四分位数は　$Q_2=\dfrac{24+31}{2}=27.5$
下半分 5 個のデータの中央値すなわち，
第 1 四分位数は　$Q_1=20$
上半分 5 個のデータの中央値すなわち，
第 3 四分位数は　$Q_3=44$

(3)　データは 11 個であるから，中央値すなわち，
第 2 四分位数は　$Q_2=33$
下半分 5 個のデータの中央値すなわち，
第 1 四分位数は　$Q_1=18$
上半分 5 個のデータの中央値すなわち，
第 3 四分位数は　$Q_3=48$

(4)　データは 13 個であるから，中央値すなわち，
第 2 四分位数は　$Q_2=41$
下半分 6 個のデータの中央値すなわち，
第 1 四分位数は　$Q_1=\dfrac{12+14}{2}=13$
上半分 6 個のデータの中央値すなわち，
第 3 四分位数は　$Q_3=\dfrac{52+54}{2}=53$

19 次の表は，変量 x，y のデータである。

x	27	37	57	59	64	72	74	80	83	84	89	94	98
y	23	40	60	66	67	73	74	78	91	96	96	97	99

(1)　x，y のデータの四分位範囲と四分位偏差を求めよ。
(2)　x，y のデータについて，四分位範囲を比較し，データの散らばり具合の大きい方をいえ。

解答　(1)　x の四分位範囲 28.5，四分位偏差 14.25
y の四分位範囲 33，四分位偏差 16.5
(2)　y

解説

(1)　[1]　変量 x のデータは 13 個であるから，下半分 6 個のデータの中央値すなわち，
第 1 四分位数は　$Q_1=\dfrac{57+59}{2}=58$
上半分 6 個のデータの中央値すなわち，
第 3 四分位数は　$Q_3=\dfrac{84+89}{2}=86.5$
四分位範囲は　$Q_3-Q_1=86.5-58=28.5$
四分位偏差は　$\dfrac{Q_3-Q_1}{2}=\dfrac{28.5}{2}=14.25$
[2]　変量 y のデータは 13 個であるから，下半分 6 個のデータの中央値すなわち，
第 1 四分位数は　$Q_1=\dfrac{60+66}{2}=63$
上半分 6 個のデータの中央値すなわち，
第 3 四分位数は　$Q_3=\dfrac{96+96}{2}=96$
四分位範囲は　$Q_3-Q_1=96-63=33$
四分位偏差は　$\dfrac{Q_3-Q_1}{2}=\dfrac{33}{2}=16.5$

(2)　y の方が四分位範囲が大きいから，y の方がデータの散らばり具合が大きい。

20 次のデータは，A 市と B 市の 1 週間の降雪量を調べたものである。
A 市　8，14，7，　9，16，　3，2　(cm)
B 市　8，35，27，31，　7，14，9　(cm)
(1)　A 市と B 市のデータの四分位範囲と四分位偏差を求めよ。
(2)　A 市，B 市のデータについて，四分位範囲を比較し，データの散らばり具合の大きい方をいえ。

解答　(1)　A 市の四分位範囲 11，四分位偏差 5.5
B 市の四分位範囲 23，四分位偏差 11.5
(2)　B 市

解説

(1)　データを小さい方から並べると
A　2，3，7，8，9，14，16
B　7，8，9，14，27，31，35
[1]　A 市のデータは 7 個であるから，下半分 3 個のデータの中央値すなわち，
第 1 四分位数は　$Q_1=3$
上半分 3 個のデータの中央値すなわち，
第 3 四分位数は　$Q_3=14$
四分位範囲は　$Q_3-Q_1=14-3=11$
四分位偏差は　$\dfrac{Q_3-Q_1}{2}=\dfrac{11}{2}=5.5$
[2]　B 市のデータは 7 個であるから，下半分 3 個のデータの中央値すなわち，
第 1 四分位数は　$Q_1=8$
上半分 3 個のデータの中央値すなわち，
第 3 四分位数は　$Q_3=31$
四分位範囲は　$Q_3-Q_1=31-8=23$
四分位偏差は　$\dfrac{Q_3-Q_1}{2}=\dfrac{23}{2}=11.5$

(2)　B 市の方が四分位範囲が大きいから，B 市の方がデータの散らばり具合が大きい。