

データの分析①

【代表値】

① 次のデータは、ある生徒12人でゲームをしたときの得点である。このデータの平均値、最頻値、中央値を求めよ。

6 5 5 6 3 3 7 3 6 4 6 6 (点)
 3 3 3 4 5 5 6 6 6 6 6 7

最頻値 6

中央値 5.5

平均値 $\frac{1}{12}(9+4+10+30+7)$
 $= \frac{60}{12} = 5$

② 右のヒストグラムは、ある高校の生徒25人について、この1週間における路線バスの利用日数を調査した結果である。

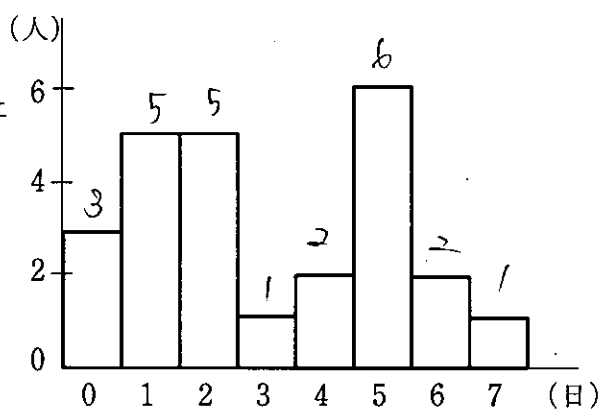
(1) 最頻値、中央値を求めよ。

最頻値 5

中央値 2

(2) 平均値を求めよ。

$\frac{1}{25}(0+5+10+3+8+30+12+7)$
 $= \frac{75}{25}$
 $= 3$



【データの散らばりと四分位数】

③ 次のデータは、ある年のA市とB市における、月ごとの降水日数を調べたものである。このデータの範囲を求めよ。また、このデータの散らばりの度合いが大きいのはA市、B市のどちらか。

A市 7, 4, 9, 7, 10, 13, 14, 7, 4, 12, 13, 5 (日)
 B市 19, 16, 12, 11, 6, 8, 21, 13, 10, 14, 18, 22 (日)

A市 $14 - 4 = 10$ (日)
 B市 $22 - 6 = 16$ (日)

$5 > 2$ B市

④ 次のデータは、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のどちらか。

A 21, 29, 32, 36, 38, 40, 49, 53, 55, 68, 80
 B 25, 31, 39, 42, 45, 46, 50, 53, 54, 65, 80

A市 四分位範囲 $55 - 32 = 23$
 四分位偏差 $\frac{23}{2} = 11.5$

B 25, 31, 39, 42, 45, 46, 50, 53, 54, 65, 80

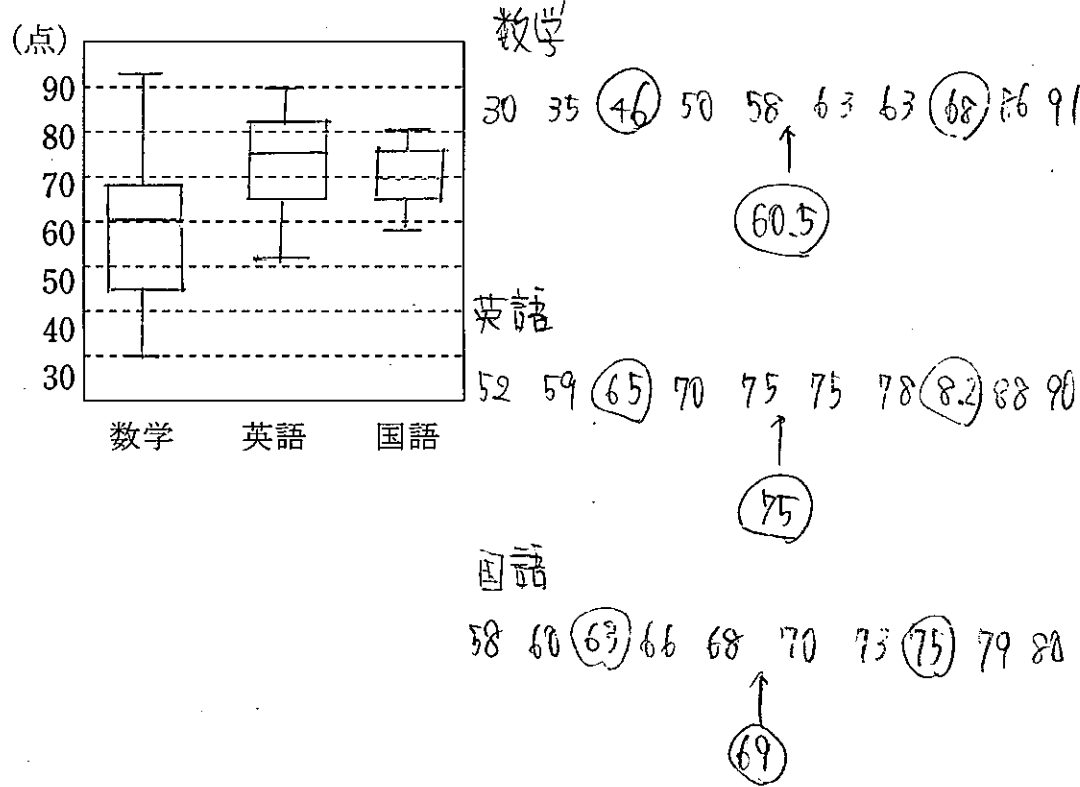
B市 四分位範囲 $54 - 39 = 15$
 四分位偏差 $\frac{15}{2} = 7.5$

$5 > 7.5$ A市

⑤ 次のデータは、10人の生徒に100点満点の数学、英語、国語のテストを行った結果である。単位は点である。

数学 68, 35, 86, 63, 30, 91, 50, 63, 46, 58
 英語 75, 65, 90, 78, 52, 88, 70, 75, 59, 82
 国語 63, 60, 73, 75, 58, 79, 68, 70, 66, 80

(1) これらのデータの箱ひげ図を並べてかけ。

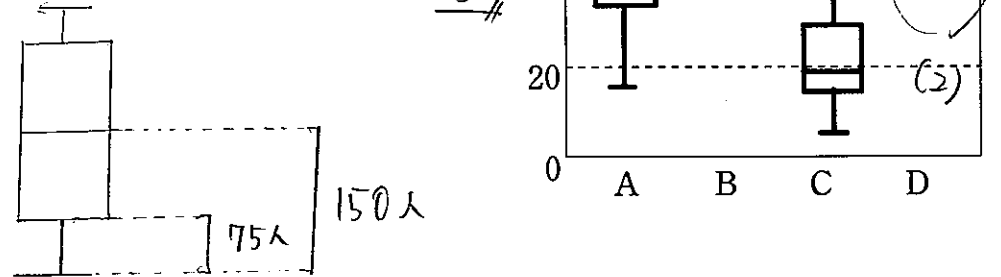


(2) データの散らばりの度合いが大きいのは、数学、英語、国語のうちどれか。

数学

⑥ 右の図は、ある学校で行った4種類のテストA、B、C、Dについての、300人の得点の箱ひげ図である。なお、どのテストも100点満点である。

- (1) 80点以上の生徒が半数以上いたのはどのテストか。
 (2) 40点以下の生徒がいて、その数が75人未満であったのはどのテストか。



【分散と標準偏差】

⑦ 次のデータは、10人の生徒に計算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。

11, 5, 12, 17, 7, 15, 9, 16, 12, 6 (点)

	11	5	12	17	7	15	9	16	12	6	110	11
$x-11$	0	-6	1	6	-4	4	-2	5	1	-5	0	0
$(x-11)^2$	0	36	1	36	16	16	4	25	1	25	160	16

分散 $\frac{60}{10} = 6$

標準偏差 $\sqrt{6} = 2.45$

データの分析②

8 次のデータについて、分散、標準偏差を求めよ。

5, 3, 6, 8, 5, 8, 5, 4, 6, 5

										計	平均	
x	5	3	6	8	5	8	5	4	6	5	55	5.5
x^2	25	9	36	64	25	64	25	16	36	25	325	32.5

2乗の平均 $\frac{325}{10} = 32.5$

平均の2乗 $(5.5)^2 = 30.25$

よって

分散 $32.5 - 30.25 = 2.25$

標準偏差 $\sqrt{2.25} = \sqrt{\frac{225}{100}} = \frac{15}{10} = \frac{3}{2} = 1.5$

9 5個の値2, 3, a, 8, 12からなるデータの平均値が6であるとき、aの値を求めよ。また、このデータの分散を求めよ。

$\frac{1}{5}(2+3+a+8+12) = 6$

$a+25 = 30$

$a = 5$

						計	平均
x	2	3	5	8	12		
$x-6$	-4	-3	-1	2	6	0	
$(x-6)^2$	16	9	1	4	36	66	13.2

分散 $\frac{66}{5} = 13.2$

10 20個の値からなるデータがあり、そのうちの8個の値の平均値は3、分散は4、残り12個の値の平均値は8、分散は9である。

(1) このデータの平均値を求めよ。

$\frac{24 + 96}{20} = \frac{120}{20} = 6$

(2) このデータの分散を求めよ。

8個のデータの2乗の和をaとおく

$4 = \frac{a}{8} - 9$

$\frac{a}{8} = 13$

$a = 104$

12個のデータの2乗の和をbとおく

$9 = \frac{b}{12} - 64$

$\frac{b}{12} = 73$

$b = 876$

20個のデータの分散は

$\frac{104 + 876}{20} - 6^2 = \frac{980}{20} - 6^2$

$= 49 - 36$

$= 13$

【データの相関】

11 次のような2つの変数x, yからなるデータがある。これらについて、散布図をかき、xとyの間に相関があるかどうかを調べよ。また、相関がある場合には、正か負のどちらの相関であるかをいえ。

(1)

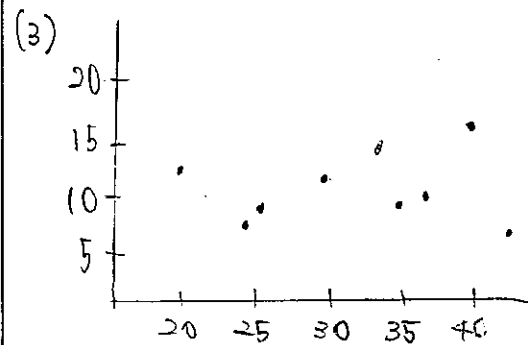
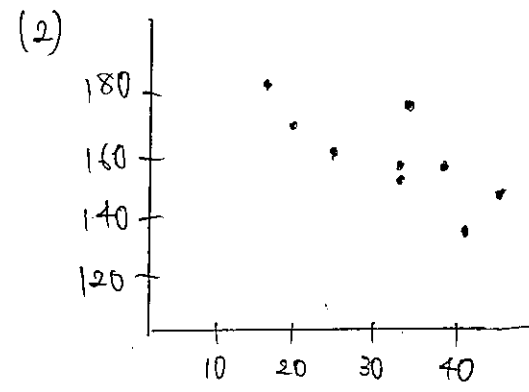
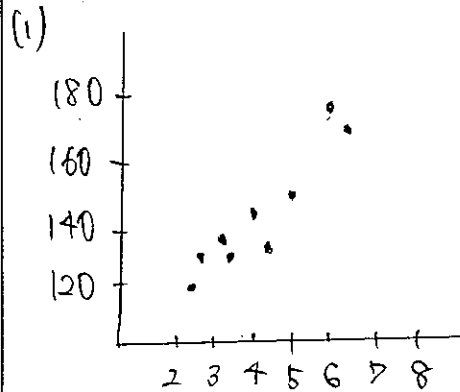
x	3.5	2.6	5.2	2.5	3.9	6.5	3.3	6.0	4.4
y	129	128	152	120	143	168	131	177	130

(2)

x	15	33	18	25	45	33	38	40	32
y	180	143	172	160	142	146	155	128	175

(3)

x	29	34	25	20	40	24	37	33	44
y	11	8	9	13	16	8	10	15	7



(1) 正の相関

(2) 負の相関

(3) 相関がない

12 下の表は、10人の生徒に10点満点の2種類のテストA, Bを行った結果である。A, Bの得点の相関係数を求めよ。また、これらの間にはどのような相関があると考えられるか。

	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩
テストA	8	9	6	2	10	3	8	4	1	9
テストB	2	2	5	5	2	5	4	4	7	4

(単位は点) 計 平均

x	8	9	6	2	10	3	8	4	1	9	60	6
$x-6$	2	3	0	-4	4	-3	2	-2	-5	3	0	
$(x-6)^2$	4	9	0	16	16	9	4	4	25	9	96	9.6
y	2	2	5	5	2	5	4	4	7	4	40	4
$y-4$	-2	-2	1	1	-2	1	0	0	3	0	0	
$(y-4)^2$	4	4	1	1	4	1	0	0	9	0	24	2.4
$(x-6)(y-4)$	-4	-6	0	-4	-8	-3	0	0	-15	0	-40	-4

よって 分散 -4

相関係数は

$\frac{-40}{\sqrt{96} \sqrt{24}} = \frac{-40}{4\sqrt{6} \cdot 2\sqrt{6}} = \frac{-40}{8 \cdot 6} = -\frac{40}{48} = -\frac{5}{6}$

$\frac{0.83}{20} \approx 0.83$

よって 負の相関