第４問 (配点　20)

以下の問題を解答するにあたっては，与えられたデータに対して，次の値を外れ値とする。

「(第１四分位数)－1.5×(四分位範囲)」以下の値
「(第３四分位数)＋1.5×(四分位範囲)」以上の値

太郎さんは，47 都道府県における外国人宿泊者数と日本人宿泊者数の動向を調べるため，それらに関するデータを分析することにした。外国人宿泊者数を，日本国内に住所を有しない宿泊者の人数の１年間の合計とし，日本人宿泊者数を，日本国内に住所を有する宿泊者の人数の１年間の合計とする。宿泊者数に関するデータは千の位を四捨五入し，１万人単位で表したものとし，以下においては単位(万人)を省略して用いることとする。例えば，「4567890 人」は「457」とする。

なお，以下の図や表については，国土交通省の Web ページをもとに作成している。

(数学 I 第４問は次ページに続く。)

(1)

(i) 図1は，47都道府県における令和4年の外国人宿泊者数と日本人宿泊者数の散布図である。なお，散布図には原点を通り，傾きが10の直線(破線)を付加している。また，日本人宿泊者数が1000を超える都道府県の数は12である。

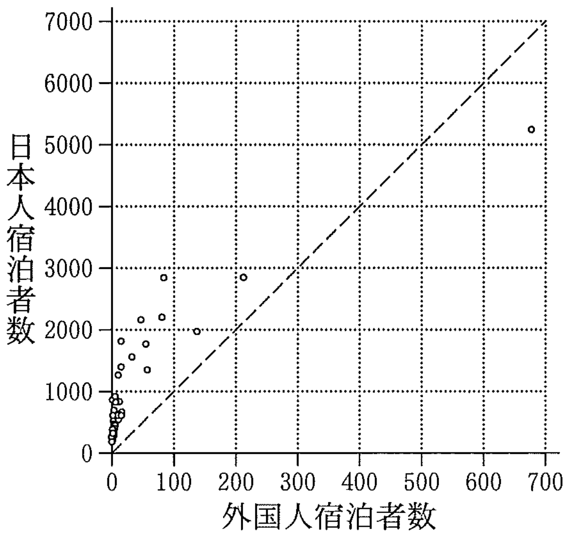


図1　令和4年の外国人宿泊者数と日本人宿泊者数の散布図

次の(a)，(b)は，図1に関する記述である。

(a) 令和4年について，外国人宿泊者数が100を超え，かつ日本人宿泊者数が2500を超える都道府県の数は2である。
(b) 令和4年について，日本人宿泊者数が外国人宿泊者数の10倍未満である都道府県の割合は50％未満である。

(a)，(b)の正誤の組合せとして正しいものは $\boxed{ア}$ である。

$\boxed{ア}$ の解答群

	⓪	①	②	③
(a)	正	正	誤	誤
(b)	正	誤	正	誤

(数学Ⅰ第4問は次ページに続く。)

(ii) 47 都道府県における令和 4 年の外国人宿泊者数を分析した結果，外れ値となる都道府県の数は 8 であった。

一方，表 1 は 47 都道府県における令和 4 年の日本人宿泊者数を，値の小さい順に並べ，その順に都道府県 P 1，P 2，…，P 47 としたものである。この中で，外国人宿泊者数で外れ値となる都道府県(P 37，P 40，P 42，P 43，P 44，P 45，P 46，P 47)に印 * を付けている。

表 1　47 都道府県における令和 4 年の日本人宿泊者数

都道府県	日本人宿泊者数	都道府県	日本人宿泊者数	都道府県	日本人宿泊者数	都道府県	日本人宿泊者数
P 1	182	P 13	373	P 25	620	P 37*	1339
P 2	187	P 14	388	P 26	625	P 38	1399
P 3	197	P 15	395	P 27	646	P 39	1547
P 4	204	P 16	401	P 28	670	P 40*	1765
P 5	255	P 17	405	P 29	683	P 41	1814
P 6	270	P 18	452	P 30	705	P 42*	1970
P 7	276	P 19	458	P 31	831	P 43*	2158
P 8	286	P 20	501	P 32	832	P 44*	2195
P 9	303	P 21	522	P 33	839	P 45*	2831
P 10	321	P 22	537	P 34	876	P 46*	2839
P 11	328	P 23	605	P 35	925	P 47*	5226
P 12	351	P 24	613	P 36	1251		

表 1 のデータにおいて，四分位範囲は $\boxed{\text{イ}}$ となることから，令和 4 年の外国人宿泊者数と日本人宿泊者数の両方で外れ値となる都道府県の数は $\boxed{\text{ウ}}$ である。

$\boxed{\text{イ}}$ の解答群

⓪	320	①	450	②	597	③	638	④	900
⑤	966	⑥	1253	⑦	1261	⑧	1602	⑨	1864

(数学 I 第 4 問は次ページに続く。)

(iii) 令和４年の外国人宿泊者数と日本人宿泊者数について，両方で外れ値となる都道府県を除いたデータで散布図を作成したところ，正の相関があることがわかった。このときの相関係数を計算するために，表２のように，平均値，標準偏差および共分散を求めた。

表２　両方で外れ値となる都道府県を除いた令和４年の外国人宿泊者数と日本人宿泊者数の平均値，標準偏差，共分散

	平均値	標準偏差	共分散
外国人宿泊者数	15	26	11373
日本人宿泊者数	739	552	

表２を用いると，両方で外れ値となる都道府県を除いた令和４年の外国人宿泊者数と日本人宿泊者数の相関係数は $\boxed{\text{エ}}$ である。

$\boxed{\text{エ}}$ については，最も適当なものを，次の⓪～⑦のうちから一つ選べ。

⓪ 0.03	① 0.21	② 0.59	③ 0.68
④ 0.79	⑤ 0.97	⑥ 1.03	⑦ 1.26

(数学Ⅰ第４問は次ページに続く。)

(2) 47 都道府県におけるある年の外国人宿泊者数を x，日本人宿泊者数を y とし，x と y の値の組を，それぞれ

$$(x_1,\ y_1),\ (x_2,\ y_2),\ \cdots,\ (x_{47},\ y_{47})$$

と表す。x，y の平均値をそれぞれ $\overline{x}$，$\overline{y}$ とし，x，y の分散をそれぞれ $s_x{}^2$，$s_y{}^2$ とする。また，x と y の共分散を s_{xy} とする。

47 都道府県それぞれにおける外国人宿泊者数と日本人宿泊者数を足し合わせた合計宿泊者数を z とし，その値を

$$z_i = x_i + y_i \quad (i = 1,\ 2,\ \cdots,\ 47)$$

と表す。例えば，$i = 7$ のときは $z_7 = x_7 + y_7$ である。

(i) z の平均値を $\overline{z}$ とするとき

$$z_i - \overline{z} = (x_i - \overline{x}) + (y_i - \overline{y}) \quad (i = 1,\ 2,\ \cdots,\ 47)$$

である。このことに着目すると，z の分散を $s_z{}^2$ とするとき，$s_z{}^2 = \boxed{\text{オ}}$ となる。

（数学Ⅰ第 4 問は次ページに続く。）

また，令和４年の x と y の間には正の相関があることが図１からわかる。このことから，令和４年について，$s_z{}^2$ と $s_x{}^2 + s_y{}^2$ の関係として，後の⓪～②のうち，正しいものは $\boxed{\text{カ}}$ であることがわかる。

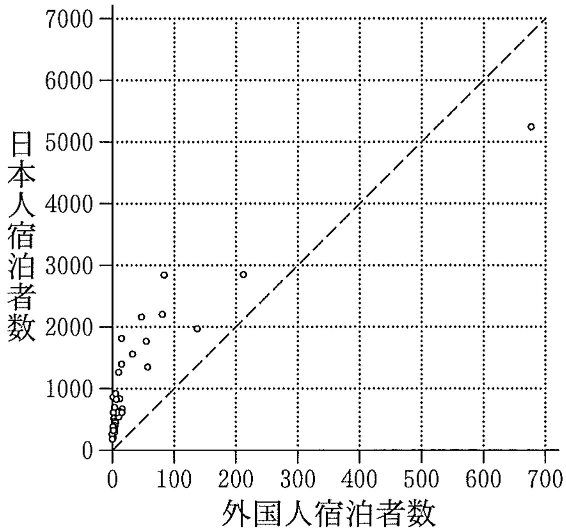


図１（再掲）

$\boxed{\text{オ}}$ の解答群

⓪ $s_x{}^2 + s_y{}^2 - 2s_{xy}$	① $s_x{}^2 + s_y{}^2 - s_{xy}$	② $s_x{}^2 + s_y{}^2$
③ $s_x{}^2 + s_y{}^2 + s_{xy}$	④ $s_x{}^2 + s_y{}^2 + 2s_{xy}$	

$\boxed{\text{カ}}$ の解答群

⓪ $s_z{}^2 > s_x{}^2 + s_y{}^2$

① $s_z{}^2 = s_x{}^2 + s_y{}^2$

② $s_z{}^2 < s_x{}^2 + s_y{}^2$

（数学Ⅰ第４問は次ページに続く。）

(ii) 太郎さんは，合計宿泊者数 z の変化に関心をもち，z についての前年との比に着目することにした。

例えば，(1)の(ii)における都道府県 P 22 の z は令和 2 年では 450 であり，令和 3 年では 376 であった。このとき，都道府県 P 22 における令和 3 年の z についての前年との比は，376 を 450 で割った値である。以下においては，それぞれの都道府県におけるある年の z についての**前年比**を，次のように定める。

前年比

ある年の z を，その前年の z で割った値

図 2 は，47 都道府県における令和元年から令和 4 年までの**前年比**の箱ひげ図を並べたものである。図 2 にある四つの箱ひげ図において，**前年比**の外れ値は，白丸で示されている。

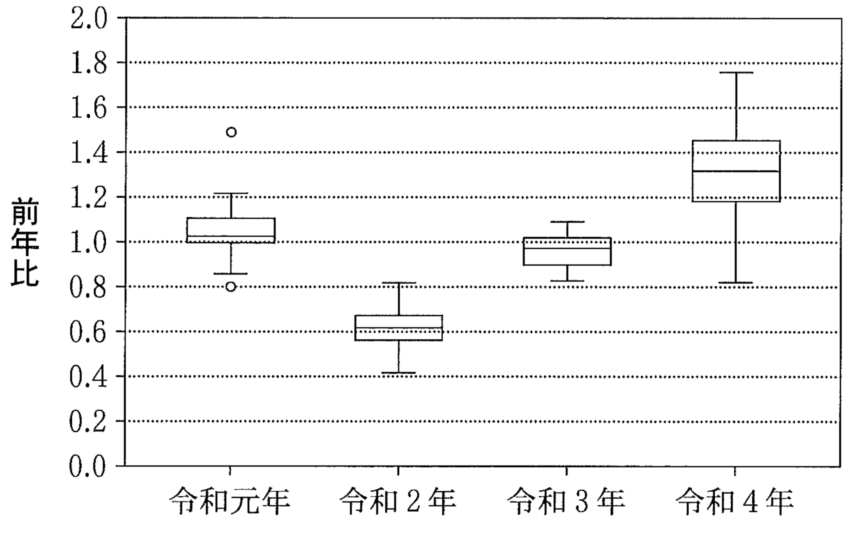


図 2　47 都道府県における令和元年から令和 4 年までの**前年比**の箱ひげ図

(数学Ⅰ第 4 問は次ページに続く。)

次の(a), (b), (c)は，図 2 に関する記述である。

(a) 令和元年の箱ひげ図において外れ値となる都道府県はすべて，令和 4 年においても外れ値となっている。

(b) すべての都道府県において，令和 2 年の z は令和元年よりも減少している。

(c) 令和 4 年の**前年比**が 1 より小さい都道府県の数は，令和 3 年の**前年比**が 1 より小さい都道府県の数よりも少ない。

(a), (b), (c)の正誤の組合せとして正しいものは $\boxed{キ}$ である。

$\boxed{キ}$ の解答群

	⓪	①	②	③	④	⑤	⑥	⑦
(a)	正	正	正	正	誤	誤	誤	誤
(b)	正	正	誤	誤	正	正	誤	誤
(c)	正	誤	正	誤	正	誤	正	誤

(数学 I 第 4 問は次ページに続く。)

(3) 太郎さんが住む地域では，その地域に宿泊を促すためのキャンペーンとして，キャンペーンA，Bが実施されている。

太郎さんは，キャンペーンAの方がよいと思っている人が多いといううわさを聞いた。このうわさのとおり，キャンペーンAの方がよいと思っている人が多いといえるかどうかを確かめることにした。そこで，かたよりなく選んだ人たちに，キャンペーンA，Bのどちらがよいかについて，二択のアンケートを行ったところ，アンケートに回答した35人のうち，23人が「キャンペーンAの方がよい」と答えた。この結果から，一般にキャンペーンAの方がよいと思っている人が多いといえるかどうかを，次の**方針**で考えることにした。

方針

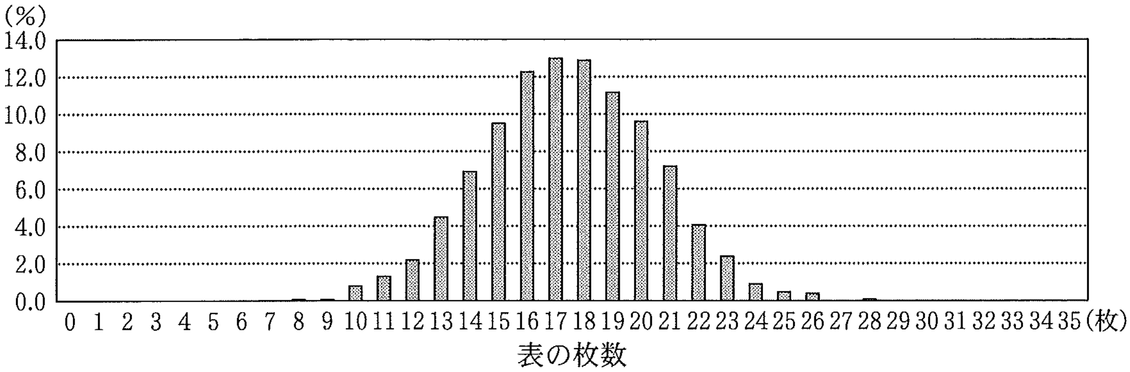
- “「キャンペーンAの方がよい」と回答する割合と「キャンペーンBの方がよい」と回答する割合は等しい”という仮説を立てる。
- この仮説のもとで，かたよりなく選ばれた35人のうち23人以上が「キャンペーンAの方がよい」と回答する確率が5％未満であれば，その仮説は誤っていると判断し，5％以上であればその仮説は誤っているとは判断しない。

後の**実験結果**は，35枚の硬貨を投げる実験を1000回行ったとき，表が出た枚数ごとの回数の割合を示したものである。

(数学Ⅰ第4問は次ページに続く。)

実験結果

表の枚数(枚)	0	1	2	3	4	5	6	7	8	9	10	11
割合(%)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.8	1.3
表の枚数(枚)	12	13	14	15	16	17	18	19	20	21	22	23
割合(%)	2.2	4.5	6.9	9.5	12.3	13.0	12.9	11.2	9.6	7.2	4.1	2.4
表の枚数(枚)	24	25	26	27	28	29	30	31	32	33	34	35
割合(%)	0.9	0.5	0.4	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0



実験結果を用いると，35 枚の硬貨のうち 23 枚以上が表となった割合は，$\boxed{\text{ク}}$. $\boxed{\text{ケ}}$ %である。これを，35 人のうち 23 人以上が「キャンペーン A の方がよい」と回答する確率とみなし，**方針**に従うと，“「キャンペーン A の方がよい」と回答する割合と「キャンペーン B の方がよい」と回答する割合は等しい” という仮説は $\boxed{\text{コ}}$ 。したがって，今回のアンケート結果からは，キャンペーン A の方がよいと思っている人が $\boxed{\text{サ}}$ 。

$\boxed{\text{コ}}$ ， $\boxed{\text{サ}}$ については，最も適当なものを，次のそれぞれの解答群から一つずつ選べ。

$\boxed{\text{コ}}$ の解答群

⓪ 誤っていると判断する	① 誤っているとは判断しない

$\boxed{\text{サ}}$ の解答群

⓪ 多いといえる	① 多いとはいえない