数学Ⅰ・数学A

第2問 （必答問題）（配点　30）

（15点）

〔1〕　太郎さんは，総務省が公表している2020年の家計調査の結果を用いて，地域による食文化の違いについて考えている。家計調査における調査地点は，都道府県庁所在市および政令指定都市（都道府県庁所在市を除く）であり，合計52市である。家計調査の結果の中でも，スーパーマーケットなどで販売されている調理食品の「二人以上の世帯の1世帯当たり年間支出金額（以下，支出金額，単位は円）」を分析することにした。以下においては，52市の調理食品の支出金額をデータとして用いる。

太郎さんは調理食品として，最初にうなぎのかば焼き（以下，かば焼き）に着目し，図1のように52市におけるかば焼きの支出金額のヒストグラムを作成した。ただし，ヒストグラムの各階級の区間は，左側の数値を含み，右側の数値を含まない。

なお，以下の図や表については，総務省のWebページをもとに作成している。

下の□の数字は小さい金額からの累計度数

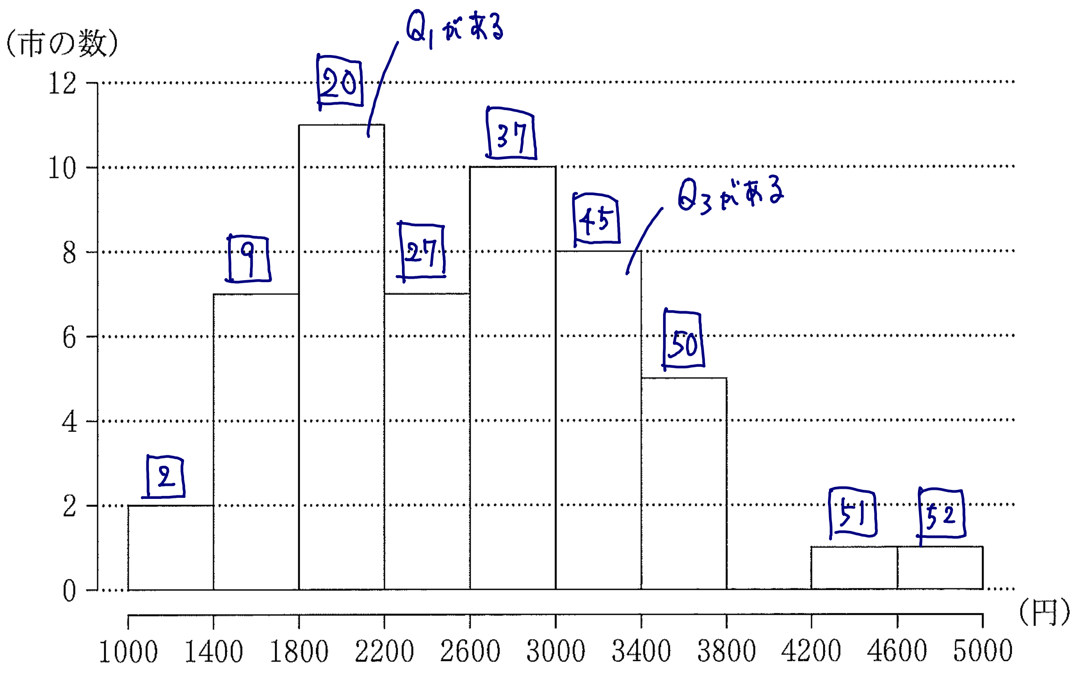


図1　かば焼きの支出金額のヒストグラム

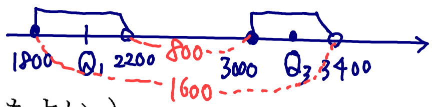
（数学Ⅰ・数学A第2問は次ページに続く。）

小さい金額から順に $x_1, x_2, \dots, x_{52}$ とすると

$x_1, x_2, \dots, x_{13} \mid x_{14}, \dots, x_{26} \mid x_{27}, \dots, x_{39} \mid x_{40}, \dots, x_{52}$

$Q_1 = \frac{x_{13}+x_{14}}{2}$ は $1800 \leqq Q_1 < 2200$ $\boxed{②}$ ア

$Q_3 = \frac{x_{39}+x_{40}}{2}$ は $3000 \leqq Q_3 < 3400$ $\boxed{⑤}$ イ



$800 < Q_3 - Q_1 < 1600$

$\boxed{①}$ ウ

(1) 図1から次のことが読み取れる。

Q_1

• 第1四分位数が含まれる階級は $\boxed{②}$ である。
(2点)

Q_3

• 第3四分位数が含まれる階級は $\boxed{⑤}$ である。
(2点)

• 四分位範囲は $\boxed{①}$。
(2点)

$Q_3 - Q_1$

$\boxed{ア}$, $\boxed{イ}$ の解答群(同じものを繰り返し選んでもよい。)

⓪ 1000 以上 1400 未満	① 1400 以上 1800 未満
ア ② 1800 以上 2200 未満	③ 2200 以上 2600 未満
④ 2600 以上 3000 未満	イ ⑤ 3000 以上 3400 未満
⑥ 3400 以上 3800 未満	⑦ 3800 以上 4200 未満
⑧ 4200 以上 4600 未満	⑨ 4600 以上 5000 未満

$\boxed{ウ}$ の解答群

⓪ 800 より小さい
① 800 より大きく 1600 より小さい
② 1600 より大きく 2400 より小さい
③ 2400 より大きく 3200 より小さい
④ 3200 より大きく 4000 より小さい
⑤ 4000 より大きい

(数学Ⅰ・数学A第2問は次ページに続く。)

(2) 太郎さんは，東西での地域による食文化の違いを調べるために，52 市を東側の地域 E(19 市)と西側の地域 W(33 市)の二つに分けて考えることにした。

EAST　　WEST

(i) 地域 E と地域 W について，かば焼きの支出金額の箱ひげ図を，図 2，図 3 のようにそれぞれ作成した。

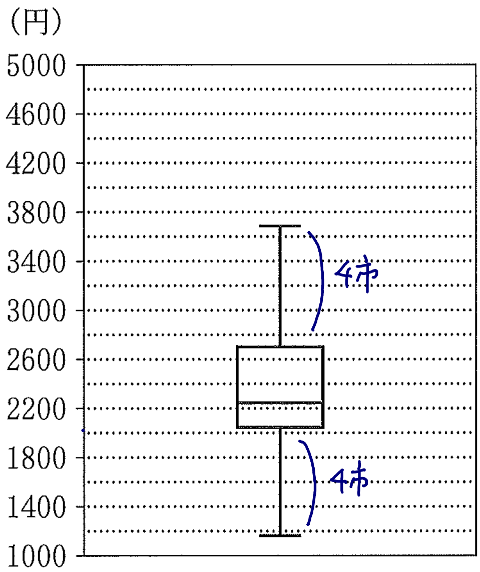


図 2　地域 E におけるかば焼きの支出金額の箱ひげ図

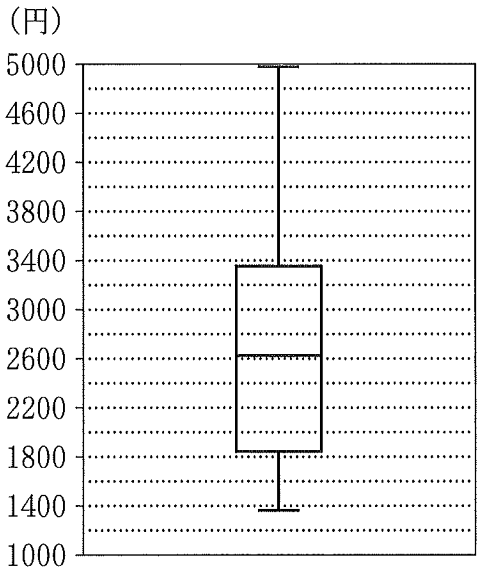


図 3　地域 W におけるかば焼きの支出金額の箱ひげ図

かば焼きの支出金額について，図 2 と図 3 から読み取れることとして，次の⓪～③のうち，正しいものは $\boxed{\text{エ}}$ ② である。

1個だけ　3個は正しくない　(3点)

E(19市)
①②③④⑤⑥⑦⑧⑨⑩⑪---⑲
Q1　Q2

$\boxed{\text{エ}}$ の解答群

⓪ 地域 E において，小さい方から 5 番目は 2000 以下である。← 箱の下線は 2000より大きい　第1四分位数

① 地域 E と地域 W の範囲は等しい。← 幅から W の方が大きい

② 中央値は，地域 E より地域 W の方が大きい。← 箱の中線から 中央値は Wの方が大きい

③ 2600 未満の市の割合は，地域 E より地域 W の方が大きい。← E は 約75%，W は 約50% で E の方が大きい

(数学Ⅰ・数学A第 2 問は次ページに続く。)

② エ

(ii) 太郎さんは，地域Eと地域Wのデータの散らばりの度合いを数値でとらえようと思い，それぞれの分散を考えることにした。地域Eにおけるかば焼きの支出金額の分散は，地域Eのそれぞれの市におけるかば焼きの支出金額の偏差の $\boxed{\text{②}}$ である。

(3点)

Eのデータを $x_1, x_2, \cdots, x_{19}$ とし 平均値を $\bar{x}$ とすると
偏差は $x_k-\bar{x}$ $(k=1, 2, \cdots, 19)$，分散は

$$\frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\cdots+(x_{19}-\bar{x})^2}{19}$$

$\boxed{②}$ オ

$\boxed{\text{オ}}$ の解答群

⓪ 2乗を合計した値
① 絶対値を合計した値
② 2乗を合計して地域Eの市の数で割った値
③ 絶対値を合計して地域Eの市の数で割った値
④ 2乗を合計して地域Eの市の数で割った値の平方根のうち正のもの ← 標準偏差
⑤ 絶対値を合計して地域Eの市の数で割った値の平方根のうち正のもの

(数学Ⅰ・数学A第2問は次ページに続く。)

↑
分散の定義を答えるだけ

(3)　太郎さんは，(2)で考えた地域Eにおける，やきとりの支出金額についても調べることにした。

ここでは地域Eにおいて，やきとりの支出金額が増加すれば，かば焼きの支出金額も増加する傾向があるのではないかと考え，まず図4のように，地域Eにおける，やきとりとかば焼きの支出金額の散布図を作成した。そして，相関係数を計算するために，表1のように平均値，分散，標準偏差および共分散を算出した。ただし，共分散は地域Eのそれぞれの市における，やきとりの支出金額の偏差とかば焼きの支出金額の偏差との積の平均値である。

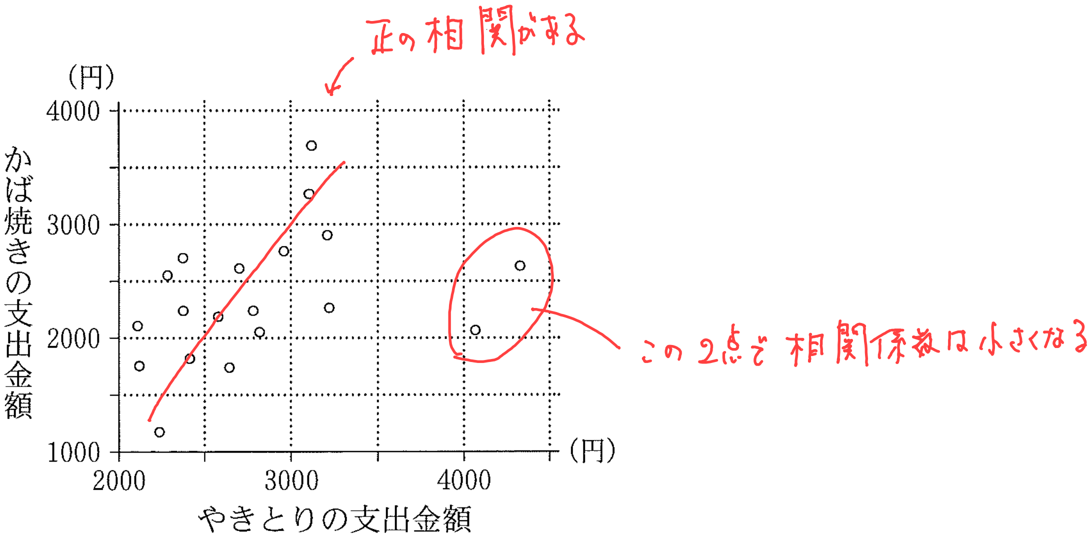


図4　地域Eにおける，やきとりとかば焼きの支出金額の散布図

表1　地域Eにおける，やきとりとかば焼きの支出金額の平均値，分散，標準偏差および共分散

	平均値	分　散	標準偏差	共分散
やきとりの支出金額	2810	348100	590	124000
かば焼きの支出金額	2350	324900	570	

(数学Ⅰ・数学A第2問は次ページに続く。)

表1を用いると，地域Eにおける，やきとりの支出金額とかば焼きの支出金額の相関係数は $\boxed{\text{カ}}$ 7 である。

(3点)

$\boxed{\text{カ}}$ については，最も適当なものを，次の⓪～⑨のうちから一つ選べ。

⓪ − 0.62	① − 0.50	② − 0.37	③ − 0.19
④ − 0.02	⑤ 0.02	⑥ 0.19	⑦ 0.37
⑧ 0.50	⑨ 0.62		

（数学Ⅰ・数学A第2問は次ページに続く。）

求める相関係数は

$$\frac{(\text{共分散})}{(\text{標準偏差の積})} = \frac{124000}{590\cdot570} = \frac{1240}{3363} = 0.368\cdots \fallingdotseq 0.37$$

$\boxed{7}$ // カ

この式でだいたい7かな

$$\begin{array}{r} 59 \\ \times\ 57 \\ \hline 413 \\ 295\ \ \\ \hline 3363 \end{array}$$

$$\begin{array}{r|l} & 0.368\cdots \\ 3363 & 12400 \\ & 10089 \\ \hline & 23110 \\ & 20178 \\ \hline & 29320 \\ & 26904 \\ \hline & 2416 \end{array}$$

補 1240 と 3363 は互いに素