

確率統計II講義ノート (補足)

Yasushi Ishikawa 2019

問1 次の図は、平成22年の都道府県別の人口10万人あたりの交通事故死亡者数についての幹葉図である。

1 | 7
2 | 0378
3 | 024444679
4 | 333445578
5 | 012233455556667
6 | 456889
7 | 123

注：人口10万人あたりの交通事故死亡者数，1.7人，2.0人，2.3人，2.7人，2.8人の死亡者数を次のように表している。

1 | 7
2 | 0378

次の記述I～IIIは、この幹葉図についての説明である。

- I. 死亡者数の平均は5.5人である。
- II. 死亡者数の中央値は5.0人である。
- III. 死亡者数の四分位範囲は約2.0人である。

これらの説明の正誤の組合せとして、適切なものを次の①～⑤のうちから一つ選べ。 1

- ① Iのみが正しい。
- ② IIのみが正しい。
- ③ IIIのみが正しい。
- ④ IとIIが正しい。
- ⑤ IIとIIIが正しい。

問2 単位がメートルで記録されている身長をセンチメートルに変更するとき、各基本統計量の数値がどう変わるかに関して、次の記述I~IVがある。

- I. 平均は100倍になる。
- II. 分散は100倍になる。
- III. 身長の変動係数は10分の1になる。
- IV. 身長と体重(キログラム)との相関係数は変化しない。

これらの記述の正誤の組合せとして、適切なものを次の①~⑤のうちから一つ選べ。 2

- ① Iのみが正しい。
- ② IとIIが正しい。
- ③ IとIIとIIIが正しい。
- ④ IとIVが正しい。
- ⑤ I,II,III,IVの全てが正しい。

問3 あるアンケート調査で、いくつかの質問をした。問1は次の形式である。

あなたは「XXX」の意見に賛成ですか、それとも反対ですか?次の1から5のうち、あてはまるもの一つを選んでその数字を○で囲んでください。

1. 賛成 2. やや賛成 3. どちらでもない 4. やや反対 5. 反対

問2は「YYY」の意見についての同様な5肢選択の質問である。

回答について、○で囲まれた数字コードを数値として記録した。問1について集計した結果は、平均が2.6、標準偏差が0.84、問2との相関係数は0.46であった。ところが、問1の質問の趣旨を踏まえれば、選択肢の数字コードを逆向きにした方がよいのではとの意見が出されたので、

「賛成」を5、「やや賛成」を4、「どちらでもない」を3、「やや反対」を2、「反対」を1

に振り直した。この数値コードを新コードと呼ぶ。新コードに基づいた平均、標準偏差、相関係数がどうなるかについて、次の記述I~IIIがある。

- I. 平均は変らない。
- II. 標準偏差は変らない。
- III. 相関係数は変らない。

これらの記述の正誤の組合せとして、適切なものを次の①~⑤のうちから一つ選べ。 3

- ① IIのみが正しい。
- ② IIIのみが正しい。
- ③ IとIIが正しい。
- ④ IIとIIIが正しい。
- ⑤ I,II,IIIの全てが間違っている。

表3は、年間収入階級別の平均年間収入である。

表3: 年間収入階級別の平均年間収入 (万円)

年間収入階級	2007年	2009年
200万円未満	161	160
200～250万円	224	226
250～300万円	275	275
300～350万円	323	323
350～400万円	372	374
400～450万円	421	423
450～500万円	473	473
500～550万円	523	521
550～600万円	572	573
600～650万円	621	621
650～700万円	673	673
700～750万円	721	721
750～800万円	773	772
800～900万円	846	843
900～1,000万円	942	945
1,000～1,250万円	1,108	1,098
1,250～1,500万円	1,361	1,356
1,500万円以上	2,106	1,971

[4] 表1, 表2, 表3から、2007年と2009年の年間収入の分布を比較するためにはどのような手法が適切か。最も適切な記述を次の①～⑤のうちから一つ選べ。

7

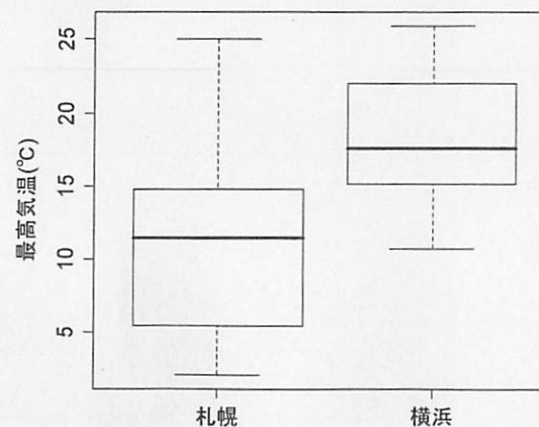
- ① 表1から求められる2007年と2009年の五数要約値を用いて検討する。
- ② 表2の平均を用いて、平均の差-13万円を643万円で割った $\frac{630-643}{643}$ を用いて検討する。
- ③ 表3の年間収入階級別の平均年間収入から2007年と2009年の両年について求めた2007年の分散と2009年の分散の分散比 $\frac{s_{2009}^2}{s_{2007}^2}$ が1に近いかどうかによって検討する。
- ④ 表1の相対度数を用いてQ-Qプロットを作成し、その傾きが1に近いかどうかで検討する。
- ⑤ 表2の2007年の年間収入の平均と2009年の年間収入の平均の差についてt検定を行う。

問5 全国の有権者を対象にした単純無作為抽出による標本調査で、内閣支持率を区間推定したい。信頼係数95%の信頼区間の幅が±4%以下となるようにするには、何人以上の有権者を調査すればよいか。最も近い数字を次の①～⑤のうちから一つ選べ。ただし、調査された人は必ず回答すると仮定する。 8

- ① 300人
- ② 420人
- ③ 600人
- ④ 840人
- ⑤ 2400人

問6 下の図は、2012年4月の30日間の札幌と横浜の最高気温についての箱ひげ図である。この箱ひげ図に関して、次の記述IからIIIがある。

- I. 札幌の最高気温の中央値はおよそ15度である。
- II. 横浜では最高気温が15度を下回る日が10日以上あった。
- III. 札幌では最高気温が10度を超えた日が15日以上あった。



これらの記述の正誤の組合せとして、適切なものを次の①～⑤のうちから一つ選べ。 9

- ① Iのみが正しい。
- ② IIのみが正しい。
- ③ IIIのみが正しい。
- ④ IとIIが正しい。
- ⑤ IとIIIが正しい。

問7 ある菓子にはおまけとして全部で15種類のカードがランダムに封入されている。カードの出る確率はすべて等しいとして以下の間に答えよ。

[1] この菓子を3個購入して、3枚とも異なるカードである確率を求めよ。小数点以下第2位までの値を、次の①～⑤のうちから一つ選べ。 10

- ① 0.58 ② 0.65 ③ 0.77 ④ 0.81 ⑤ 0.87

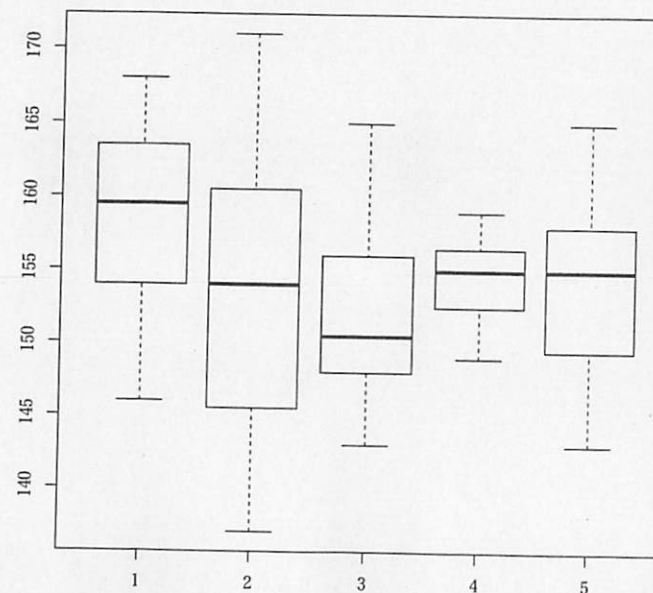
[2] 既に13種類のカードを持っているとする。まだ持っていない2種類のカードのいずれかが出るまでこの菓子を購入し続けるとすると、購入個数の期待値はいくらか。次の①～⑤のうちから一つ選べ。 11

- ① 5個 ② 7.5個 ③ 10個 ④ 12.5個 ⑤ 15個

問8 あるファストフード店におけるフライドポテトの量が公表している量よりも少ないのではないかとこの噂を検証するため、時間帯をランダムに選んでフライドポテトを20回購入して重量を計測したところ、次のような結果を得た。

平均 154.4g	中央値 155.0g	標準偏差 5.85g	四分位範囲 8.2g
-----------	------------	------------	------------

[1] このデータの箱ひげ図として最も適切なものを、次の①～⑤のうちから一つ選べ。 12

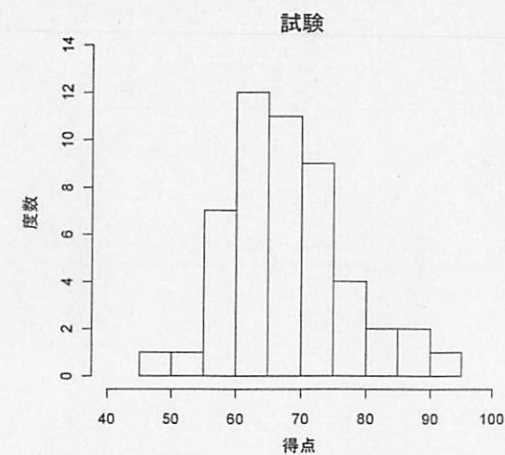


- ① 箱ひげ図1 ② 箱ひげ図2 ③ 箱ひげ図3
④ 箱ひげ図4 ⑤ 箱ひげ図5

[2] 公表されている重量は160gである。公表値よりも少ないと言えるかどうか、帰無仮説を $H_0: \mu = 160g$ として t 統計量により片側検定を行うとき、検定統計量として最も適切なものを、次の①～⑤のうちから一つ選べ。 13

- ① $\sqrt{20}(154.4 - 160)/5.85$ ② $(154.4 - 160)/(\sqrt{20} \times 5.85)$
③ $(154.4 - 160)/(5.85^2/20)$ ④ $(160 - 154.4)/(\sqrt{19}/5.85)$
⑤ $(160 - 154.4)/5.85$

問1 A先生は、担当科目について0点から100点までの1点刻みの得点となる試験を行った。次のヒストグラムは、受験した50人についての試験結果である。ただし、ヒストグラムは区間幅を5点刻み $(45, 50], (50, 55], \dots, (85, 90], (90, 95]$ として作成している。



このヒストグラムに関して、次の記述aからcがある。

- a この試験の平均は65点以上である。
- b この試験の中央値は66点から70点の区間にある。
- c 60点以下の受験者の割合はほぼ1割である。

この記述について、次の選択肢①～⑤のうちから適切なものを一つ選べ。 1

- ① aのみ正しい
- ② bのみ正しい
- ③ cのみ正しい
- ④ aとbのみ正しい
- ⑤ aとcのみ正しい

問2 B君は経済学のレポート作成で、2009年の貯蓄に関するデータを用いることにした。ただし、次の表の相対度数と累積相対度数は小数第5位で四捨五入している。

貯蓄階級	相対度数	累積相対度数
100万円未満	0.1118	0.1118
100～200万円未満	0.0640	0.1759
200～300万円未満	0.0569	0.2327
300～400万円未満	0.0547	0.2874
400～500万円未満	0.0503	0.3377
500～600万円未満	0.0476	0.3853
600～700万円未満	0.0412	0.4266
700～800万円未満	0.0358	0.4624
800～900万円未満	0.0366	0.4989
900～1000万円未満	0.0334	0.5323
1000～1200万円未満	0.0575	0.5898
1200～1400万円未満	0.0481	0.6379
1400～1600万円未満	0.0395	0.6774
1600～1800万円未満	0.0336	0.7110
1800～2000万円未満	0.0294	0.7404
2000～2500万円未満	0.0600	0.8004
2500～3000万円未満	0.0444	0.8447
3000～4000万円未満	0.0584	0.9032
4000万円以上	0.0968	1.0000

出典：総務省『家計調査』（「貯蓄・負債編」二人以上の世帯），2009年

[1] 第1四分位数が含まれる階級を、次の①～⑤のうちから一つ選べ。 2

- | | |
|---------------|---------------|
| ① 100万円未満 | ② 100～200万円未満 |
| ③ 200～300万円未満 | ④ 300～400万円未満 |
| ⑤ 400～500万円未満 | |

90パーセンタイルが含まれる階級を、次の①～⑤のうちから一つ選べ。 3

- | | |
|-----------------|-----------------|
| ① 1800～2000万円未満 | ② 2000～2500万円未満 |
| ③ 2500～3000万円未満 | ④ 3000～4000万円未満 |
| ⑤ 4000万円以上 | |

[3] この表から推測される分布として、次の①～⑤のうちから最も適切なもの一つ選べ。 4

- | | | |
|----------|----------|------|
| ① 右に裾が長い | ② 左に裾が長い | ③ 対称 |
| ④ カウ・ベル型 | ⑤ 一様 | |

[4] この分布の平均と中央値に関する記述について、次の①～⑤のうちから最も適切なもの一つ選べ。 5

- ① 平均は中央値より小さい。
- ② 平均と中央値は等しい。
- ③ 平均は中央値より大きい。
- ④ 平均を正確に求めることができる。
- ⑤ 中央値を正確に求めることができる。

問3 B君は、10点満点の4回のテストで5, 4, 7, 6の得点を取った。5回目のテストで最低何点取れば5回のテストの平均が6点以上となるか。次の①～⑤のうちから正しいもの一つ選べ。 6

- | | | |
|------|------|------|
| ① 5点 | ② 6点 | ③ 7点 |
| ④ 8点 | ⑤ 9点 | |

問13 次の表は、平成15年度から平成24年度までの10年間の、東京都の小学校に在籍する児童の総数（全児童数）と長期欠席児童のうち理由が不登校である児童の数（不登校児童数）である。

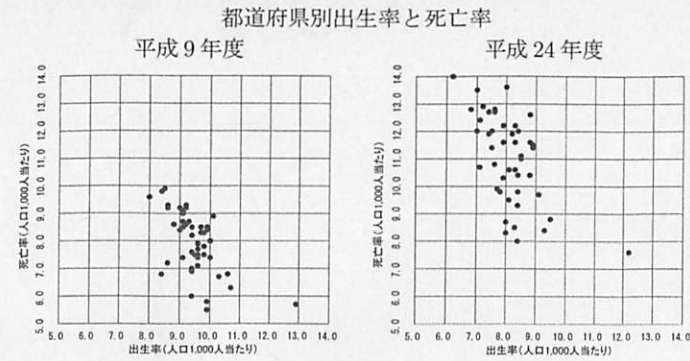
	全児童数 (人)	不登校 児童数(人)
平成15年度	567,216	1,889
平成16年度	572,575	1,832
平成17年度	580,036	1,803
平成18年度	586,492	1,907
平成19年度	588,374	1,914
平成20年度	592,736	1,862
平成21年度	594,326	1,907
平成22年度	595,669	1,969
平成23年度	592,192	2,056
平成24年度	586,412	1,940

資料：文部科学省「学校基本調査」

この10年間の不登校児童数の変化の様子を調べた結果の解釈について、次の①～⑤のうちから最も適切なものを一つ選べ。 **17**

- ① 各年度の(全児童数) - (不登校児童数)を調べてみると、ほぼ毎年増加しているので、不登校児童数は減少傾向にあるといえる。
- ② (平成24年度の不登校児童数) - (平成15年度の不登校児童数)を調べてみると、長期的には増加しているので、不登校児童の割合は増加傾向にあるといえる。
- ③ 各年度の(不登校児童数) ÷ (全児童数) × 100を調べてみると、ほぼ毎年同じ値なので、不登校児童の割合はほぼ変わらないといえる。
- ④ (平成16年度の不登校児童数) - (平成15年度の不登校児童数)、(平成17年度の不登校児童数) - (平成16年度の不登校児童数)、…、(平成24年度の不登校児童数) - (平成23年度の不登校児童数)を順次調べてみると、増加と減少の両方が見られたので、不登校児童数の変化の様子に特徴的な傾向がないといえる。
- ⑤ (平成23年度の不登校児童数)が最大であるが、(平成24年度の不登校児童数)では減少していることから、不登校児童数が減少傾向にあるといえる。

問14 次の2つの散布図は、平成9年度と平成24年度の都道府県別の1,000人当たりの出生率と死亡率を表している。

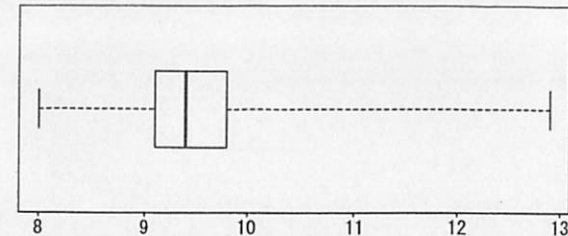


資料：厚生労働省「人口動態統計」

[1] この2つの散布図からわかることとして、適切でないものを、次の①～⑤のうちから一つ選べ。 **18**

- ① 平成9年度に比べて平成24年度の方が、全体的に死亡率が高くなっている。
- ② 平成9年度に比べて平成24年度の方が、全体的に出生率が高くなっている。
- ③ 平成24年度は、死亡率よりも出生率の方が低い都道府県が多い。
- ④ 平成9年度も平成24年度も、他の都道府県に比べてかなり出生率が高い県がみられる。
- ⑤ 平成9年度も平成24年度も、出生率と死亡率の間には負の相関がみられる。

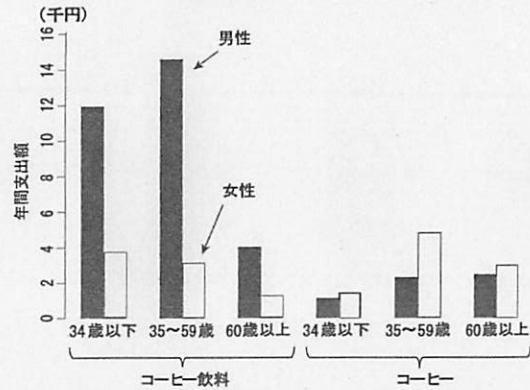
[2] 次の箱ひげ図は何を表しているのか、下の①～⑤のうちから最も適切なものを一つ選べ。 **19**



- ① 平成9年度の出生率
- ② 平成24年度の出生率
- ③ 平成9年度の死亡率
- ④ 平成24年度の死亡率
- ⑤ 平成9年度から平成24年度の間出生率の変化

問16 次の図は、単身世帯のコーヒー飲料及びコーヒーへの年間支出額（平成22～24年平均値）を男女年齢階級別にまとめたものである。家計調査では、コーヒーを主原料とした液体飲料を「コーヒー飲料」としており、レギュラーコーヒー（豆、粉など）、インスタントコーヒーなど液体以外を「コーヒー」としている。

コーヒー飲料及びコーヒーの男女年齢階級別年間支出額
（単身世帯・平成22～24年平均値）



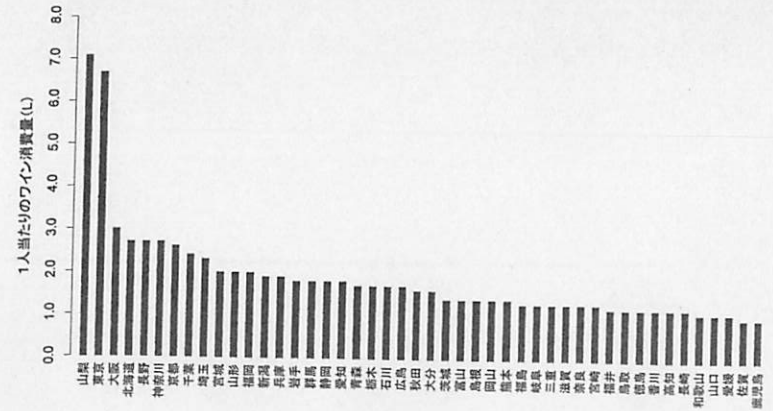
資料：家計調査通信 479号

この結果の解釈として、適切でないものを、次の①～⑤のうちから一つ選べ。

23

- ① コーヒー飲料の年間支出額は、どの年齢階級でも男性の方が多い。
- ② コーヒーの年間支出額は、どの年齢階級でも女性の方が多い。
- ③ 男性においては、60歳以上のコーヒー飲料の年間支出額は、ほかの年齢階級に比べて半分以下である。
- ④ 35～59歳の階級では、男性のコーヒー飲料の年間支出額は女性のコーヒー飲料の年間支出額の4倍以上である。
- ⑤ 男女どの年齢階級においても、コーヒー飲料の年間支出額はコーヒーの年間支出額を上回っている。

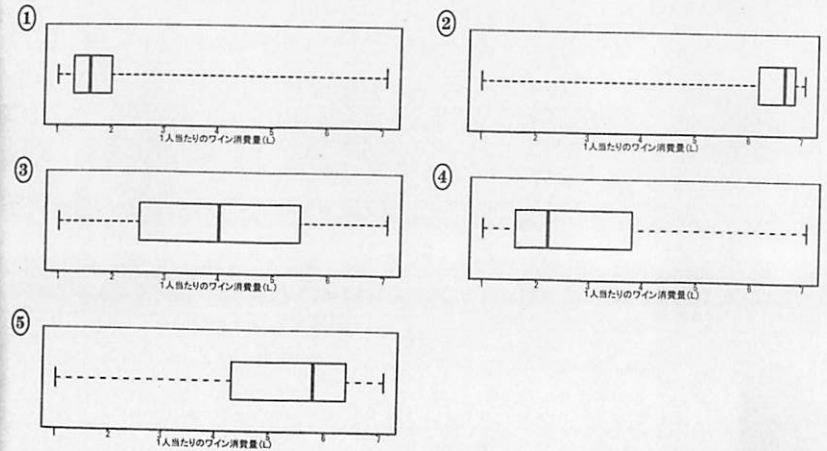
問17 次の棒グラフは、平成22年度の沖縄県を除く46都道府県の成人1人当たりのワイン消費量（L）を表している。



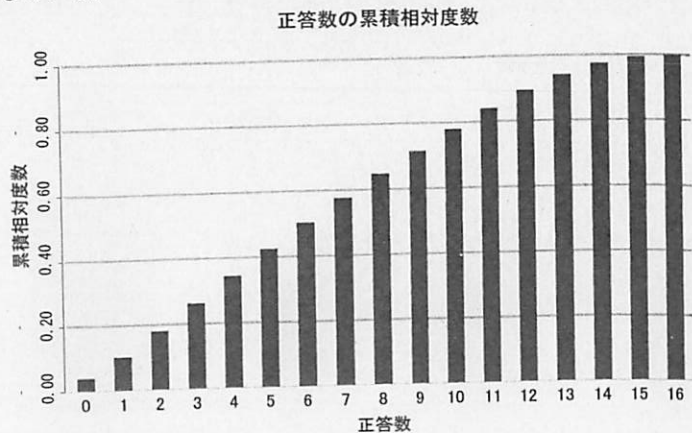
資料：国税庁課税部酒税課「平成22年度酒類販売（消費）数量表」

この結果の中で、7.1Lの山梨県が最も多く、次いで東京都の6.7Lであり、最も少ないのは鹿児島県、佐賀県で消費量は1.0Lであった。この棒グラフから作成した46の都道府県のワイン消費量の箱ひげ図として、次の①～⑤のうちから最も適切なものを一つ選べ。

24



問18 ある学校で16問のテストを行った。次の図は横軸に正答数を少ない順から0~16を1つ刻みでとり、縦軸にその数以下を正答した受験者の割合（累積相対度数）を棒状に表している。



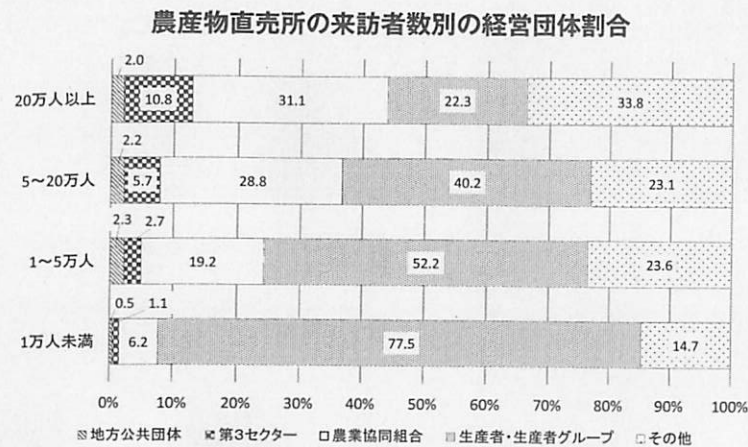
このように、累積相対度数のグラフを用いる利点として、適切でないものを、次の①~⑤のうちから一つ選べ。 25

- ① 正答数の中央値をグラフから読み取ることができる。
- ② 正答数が8問以下の受験者の割合をグラフから読み取ることができる。
- ③ 正答数が5問以上10問以下の受験者の割合をグラフから読み取ることができる。
- ④ 正答数の平均値をグラフから読み取ることができる。
- ⑤ 上位20%に入った受験者は、何問以上正答したかをグラフから読み取ることができる。

問19 1から6の目が同じ確率で出るサイコロがある。このサイコロを4回投げたときに、偶数の目が出る回数が2回以下となる確率として正しいものを、次の①~⑤のうちから一つ選べ。 26

- ① $\frac{3}{8}$
- ② $\frac{2}{5}$
- ③ $\frac{1}{2}$
- ④ $\frac{3}{5}$
- ⑤ $\frac{11}{16}$

問20 次の帯グラフは、平成21年度の農産物直売所の来訪者数と経営団体の関係を表している。



資料：農林水産省「農産物地産地消等実態調査(平成21年度)」

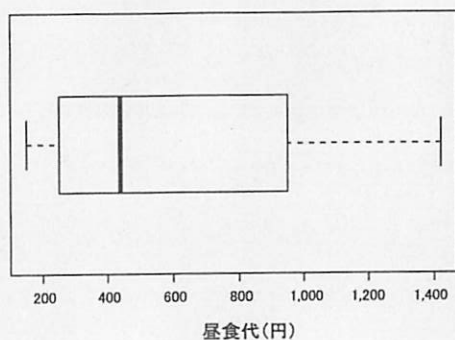
この帯グラフから読み取れることとして、次の①~⑤のうちから最も適切なものを一つ選べ。 27

- ① 来訪者が多い農産物直売所ほど、農業協同組合が経営している割合が高くなる。
- ② 農産物直売所の中では、生産者・生産者グループが経営している1万人未満のものが一番多い。
- ③ 第3セクターが経営している農産物直売所は、1万人未満のものは少ないが、1~5万人のもの、5~20万人のもの、20万人以上のものはほぼ同じ数である。
- ④ 農業協同組合が経営している農産物直売所は、20万人以上のものが一番多い。
- ⑤ 20万人以上の農産物直売所が最も数が多い。

問3 AチームとBチームのサッカーの試合において、じゃんけんで勝った方を先攻とし、あいこの場合はAチームを先攻と決めた。このとき、3回の試合の先攻を決める場合にあいこが1度も起きず、Bチームが少なくとも一度は先攻になる確率を、次の①～⑤のうちから一つ選べ。 3

- ① $1 - \frac{1}{3}$ ② $1 - \left(\frac{1}{3}\right)^3$ ③ $3 \times \left(\frac{1}{3}\right)^3$ ④ $7 \times \left(\frac{1}{3}\right)^3$ ⑤ $8 \times \left(\frac{1}{3}\right)^3$

問4 大学生200人を対象にある日の昼食にいくら支払ったかについて調査し、得られた結果を次のように箱ひげ図としてまとめた。



この箱ひげ図に対応する度数分布表として最も適切なものを、次の①～⑤のうちから一つ選べ。 4

①

昼食代	度数	累積度数
100円以上 200円未満	12	12
200円以上 300円未満	29	41
300円以上 400円未満	36	77
400円以上 500円未満	34	111
500円以上 600円未満	21	132
600円以上 700円未満	9	141
700円以上 800円未満	3	144
800円以上 900円未満	4	148
900円以上 1000円未満	19	167
1000円以上 1100円未満	13	180
1100円以上 1200円未満	5	185
1200円以上 1300円未満	0	185
1300円以上 1400円未満	11	196
1400円以上 1500円未満	4	200

②

昼食代	度数	累積度数
100円以上 200円未満	4	4
200円以上 300円未満	13	17
300円以上 400円未満	36	53
400円以上 500円未満	33	86
500円以上 600円未満	32	118
600円以上 700円未満	22	140
700円以上 800円未満	17	157
800円以上 900円未満	0	157
900円以上 1000円未満	15	172
1000円以上 1100円未満	0	172
1100円以上 1200円未満	0	172
1200円以上 1300円未満	17	189
1300円以上 1400円未満	7	196
1400円以上 1500円未満	4	200

③

昼食代	度数	累積度数
100円以上 200円未満	20	20
200円以上 300円未満	32	52
300円以上 400円未満	46	98
400円以上 500円未満	23	121
500円以上 600円未満	14	135
600円以上 700円未満	5	140
700円以上 800円未満	0	140
800円以上 900円未満	12	152
900円以上 1000円未満	31	183
1000円以上 1100円未満	9	192
1100円以上 1200円未満	1	193
1200円以上 1300円未満	0	193
1300円以上 1400円未満	0	193
1400円以上 1500円未満	7	200

④

昼食代	度数	累積度数
100円以上 200円未満	40	40
200円以上 300円未満	47	87
300円以上 400円未満	32	119
400円以上 500円未満	28	147
500円以上 600円未満	9	156
600円以上 700円未満	4	160
700円以上 800円未満	6	166
800円以上 900円未満	0	166
900円以上 1000円未満	9	175
1000円以上 1100円未満	13	188
1100円以上 1200円未満	5	193
1200円以上 1300円未満	4	197
1300円以上 1400円未満	0	197
1400円以上 1500円未満	3	200

⑤

昼食代	度数	累積度数
100円以上 200円未満	12	12
200円以上 300円未満	43	55
300円以上 400円未満	36	91
400円以上 500円未満	28	119
500円以上 600円未満	9	128
600円以上 700円未満	4	132
700円以上 800円未満	15	147
800円以上 900円未満	0	147
900円以上 1000円未満	16	163
1000円以上 1100円未満	13	176
1100円以上 1200円未満	5	181
1200円以上 1300円未満	4	185
1300円以上 1400円未満	11	196
1400円以上 1500円未満	4	200

問5 あるクラスで睡眠と成績の関係について話し合い、意見 A「極度に眠るときは一度寝てから朝早起きして試験勉強した方が点数が上がる」と意見 B「極度に眠るときでも、納得するまで勉強してから寝た方が点数が上がる」の2つがあった。そこで、次の期末試験のときに、これらの行動をとった人に対して、前回の試験との点数の変化について調査したところ、次の表ようになった。なお前回の試験の際には通常時の睡眠をとり、点数を上げるために睡眠のとり方を変える行動はしなかったと仮定する。

	前回より 上がった	前回と変わら なかった	前回より 下がった	合計
意見 A を実施した人	27	11	14	52
意見 B を実施した人	32	33	34	99

[1] 意見 A を実施した人のうち点数が上がった人の割合と意見 B を実施した人のうち点数が上がった人の割合の比として適切なものを、次の①～⑤のうちから一つ選べ。 5

- ① 27 : 32 ② $\frac{27}{52} : \frac{32}{99}$ ③ $\frac{38}{52} : \frac{65}{99}$
 ④ $\frac{27}{59} : \frac{32}{59}$ ⑤ $\frac{27}{151} : \frac{32}{151}$

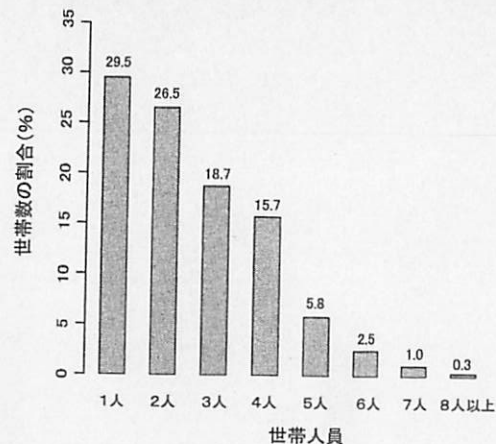
[2] この結果を踏まえ、点数を上げるための行動として、次の3つを考えた。

- I. 意見 A を実施して点数が上がった人よりも、意見 B を実施して点数が上がった人が多いため、意見 B を実施する方が点数が上がる行動と考える。
- II. 意見 B を実施した人における点数が上がった人の割合よりも、意見 A を実施した人における点数が上がった人の割合が大きいため、意見 A を実施する方が点数が上がる行動と考える。
- III. 意見 B を実施した人の「前回より下がった人」に対する「前回より上がった人」の比の値が1よりも小さいため、意見 B の実施は点数が上がる行動とは考えにくい。

この I～III の考えについて、最も適切なものを、次の①～⑤のうちから一つ選べ。 6

- ① I のみ正しい。 ② II のみ正しい。
 ③ I と II のみ正しい。 ④ I と III のみ正しい。
 ⑤ II と III のみ正しい。

問6 次の棒グラフは、平成 17 年における世帯人員別一般世帯数の割合 (%) を示したものである。



資料：総務省「平成 17 年国勢調査」

[1] 世帯人員の中央値を、次の①～⑤のうちから一つ選べ。 7

- ① 1人 ② 2人 ③ 3人 ④ 4人 ⑤ 5人

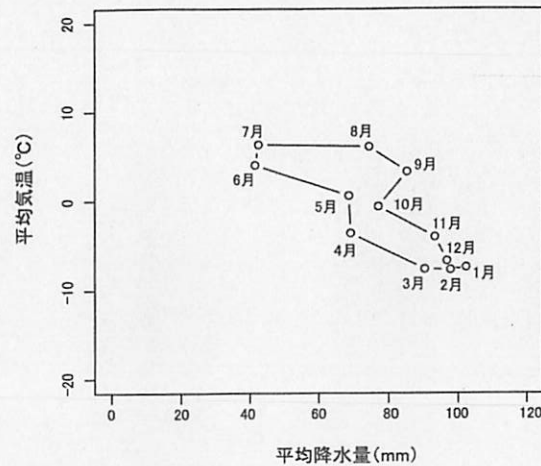
[2] 世帯全体における世帯人員が5人以上の世帯の割合として最も適切なものを、次の①～⑤のうちから一つ選べ。 8

- ① 約 4% ② 約 6% ③ 約 8% ④ 約 10% ⑤ 約 12%

[3] 上の棒グラフから平均値を求める記述として最も適切なものを、次の①～⑤のうちから一つ選べ。ただし、ここでは8人以上は8人として計算する。 9

- ① 平均値は、(世帯人員数 × 割合 × 0.01) の和で計算する。この場合、棒グラフによる平均値は、実際の平均値よりも大きくはならない。
- ② 平均値は、(世帯人員数 × 割合 × 0.01) の和で計算する。この場合、棒グラフによる平均値は、実際の平均値よりも小さくはならない。
- ③ 平均値は、(世帯人員数 × 割合 × 0.01) の和で計算する。この場合、棒グラフによる平均値は、実際の平均値と必ず一致する。
- ④ 平均値は、(世帯人員数 × 割合 × 0.01) の和を階級の数8で割る。この場合、棒グラフによる平均値は、実際の平均値よりも大きくはならない。
- ⑤ 平均値は、(世帯人員数 × 割合 × 0.01) の和を階級の数8で割る。この場合、棒グラフによる平均値は、実際の平均値よりも小さくはならない。

問7 次のグラフは、ある都市の月ごとの5年間の平均降水量（月合計降水量の5年間の平均値）と平均気温（月平均気温の5年間の平均値）を示したものである。横軸に平均降水量（mm）、縦軸に平均気温（℃）をとり、各月の記録を描き、1月から12月までを順に線で結んでいる。



このグラフの解釈として、次の3つを考えた。

- I. この都市の8月の平均気温は高く、9月、10月、…、12月と1月に近づくにつれ、平均気温が単調に低くなる傾向がある。
- II. この都市の7月の平均降水量は少なく、8月、9月、…、12月と1月に近づくにつれ、平均降水量が単調に多くなる傾向がある。
- III. この都市の5月から10月の間の平均気温はほとんど変わらず、平均降水量も10mm程度しか変化しない。

このI～IIIの解釈について、最も適切なものを、次の①～⑤のうちから一つ選べ。 10

- ① I, II, III はどれも正しい。
- ② I と II のみ正しい。
- ③ I と III のみ正しい。
- ④ II と III のみ正しい。
- ⑤ I のみ正しい。

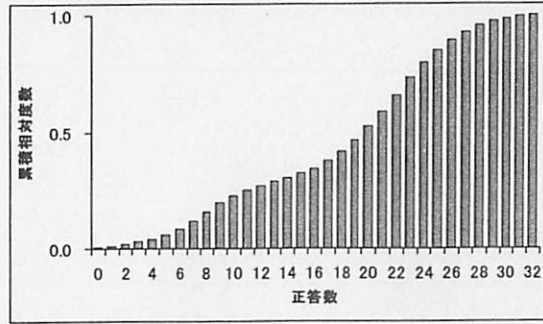
問8 ある店舗において、同じ受付作業であっても窓口によって作業時間に違いがあることが分かった。そこで各窓口での作業時間を測定して、作業時間の散らばりを調整する改善を考える。これまでの測定データから他の観測値と大きくはずれた観測値が得られることも予想され、各窓口における作業時間の分布は左右対称とはいえない。今回の調整では、はずれた観測値の影響を受けにくい指標を用いて、データの散らばりの大きさを調整することを考える。この店舗の窓口における作業時間の改善を考えるために、最も適切なものを、次の①～⑤のうちから一つ選べ。 11

- ① 各窓口の作業時間の散らばりを標準偏差で測り、値を小さくすることを考える。
- ② 各窓口の作業時間の散らばりを分散で測り、値を大きくすることを考える。
- ③ 各窓口の作業時間の散らばりを範囲で測り、値を大きくすることを考える。
- ④ 各窓口の作業時間の散らばりを四分位範囲で測り、値を小さくすることを考える。
- ⑤ 各窓口の作業時間の散らばりを変動係数で測り、値を大きくすることを考える。

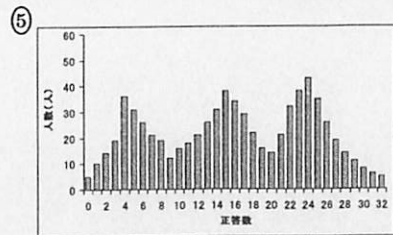
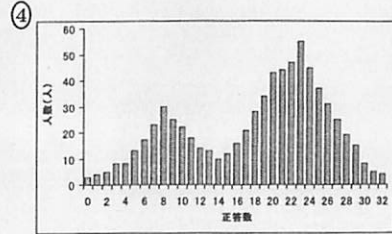
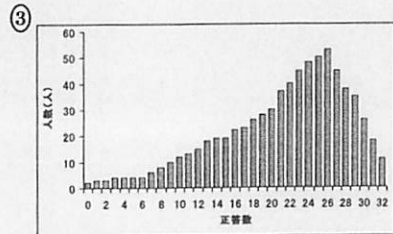
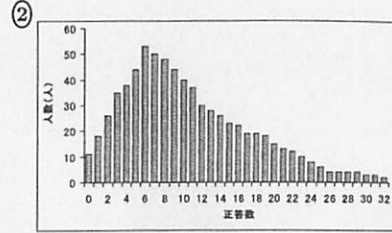
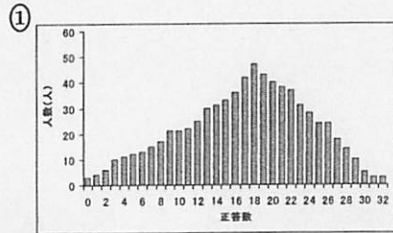
問9 総務省の「平成23年通信利用動向調査」では、全国の満20歳以上の世帯構成員がいる世帯から無作為に40,592世帯を抽出し、通信利用動向に関する調査を行っている（なお回収率は41.3%であった）。標本抽出や未回収、無回答等による標本の偏りはしないものとする。パソコンや携帯電話等から1年間にインターネットを利用したかどうかの質問に対して、71.2%の人が「利用した」と回答した。この結果を説明する文章として、最も適切なものを、次の①～⑤のうちから一つ選べ。なお、母集団におけるインターネットを「利用した」という回答の比率を母比率、標本におけるインターネットを「利用した」という回答の比率を標本比率と表現することとする。 12

- ① 標本調査により母比率を推定したい。母比率は0.712であった。
- ② 標本調査により母比率を推定したい。標本比率は0.712であった。
- ③ 標本調査により標本比率を推定したい。母比率は0.712であった。
- ④ 標本調査により標本比率を推定したい。標本比率は0.712であった。
- ⑤ 今回の調査では母比率は推定できない。

問10 ある学校で32問の数学のテストを行った。次の図は横軸に正答数を少ない順から0~32を1つ刻みでとり、縦軸にその数以下正答した生徒の割合を棒グラフで用いて表している。



このデータを用いて、正答数のヒストグラムを描いたときに得られるグラフとして適切なものを、次の①~⑤のうちから一つ選べ。 **10**



問11 行政機関等では、さまざまな統計を作成している、これを公的統計という。その中で国が作成し重要性が特に高い統計を「基幹統計」といい、この基幹統計を作成するために実施される統計調査を基幹統計調査と呼んでいる。基幹統計調査ではないものを、次の①~⑤のうちから一つ選べ。 **11**

- ① 労働力調査
- ② 家計調査
- ③ 内閣支持率調査
- ④ 社会生活基本調査
- ⑤ 国勢調査

問12 ある工場で、生産した製品が不良品である確率は2%である。また、製品の品質検査を行う際に、その製品が良品であるときに正しく良品であると判断する確率も、不良品であるときに正しく不良品であると判断する確率も99%である。このとき、不良品と判断された製品が、本当に不良品である確率として最も近いものを、次の①~⑤のうちから一つ選べ。 **12**

- ① 約2%
- ② 約50%
- ③ 約67%
- ④ 約95%
- ⑤ 約99%

問13 次の表は、あるクラスの生徒20人に対して、10点満点の数学の小テストを2回行った結果をまとめたものである。横軸は1回目の小テストの得点を、縦軸は2回目の小テストの得点を表す。また、表中の数値は人数を表し、例えば、1回目の得点が4点で2回目が6点の生徒は3人いる。ただし、空欄のセルは該当者がいないことを表す。

	10																				
	9																			1	
	8					2		1													
	7				2	1															
2	6			1	3	1	1	1													
回	5					2															
目	4					2															
	3		1			1															
	2																				
	1																				
	0																				
		0	1	2	3	4	5	6	7	8	9	10									
																					1回目

[1] 1回目よりも2回目の小テストの得点が高かった生徒の人数を、次の①～⑤のうちから一つ選べ。 13

- ① 4人 ② 5人 ③ 8人 ④ 12人 ⑤ 16人

[2] 2回の小テストの平均点を求めたところ、1回目の小テストの平均点はちょうど5点、2回目の平均点はちょうど6点であった。2回の小テストのいずれにおいても、平均点を超える得点をとった生徒の人数を、次の①～⑤のうちから一つ選べ。 14

- ① 2人 ② 3人 ③ 4人 ④ 5人 ⑤ 7人

[3] 1回目と2回目の小テストの相関係数を調べるために分散を求めてみたところ、1回目の小テストの分散は6.4、2回目の小テストの分散は5.2であった。相関係数を求める式として、適切な式を、次の①～⑤のうちから一つ選べ。 15

① $\frac{1}{\sqrt{6.4}\sqrt{5.2}} \left\{ \frac{(1-5)(3-6) + (4-5)(7-6) + (7-5)(8-6) + (10-5)(9-6)}{20} \right\}$

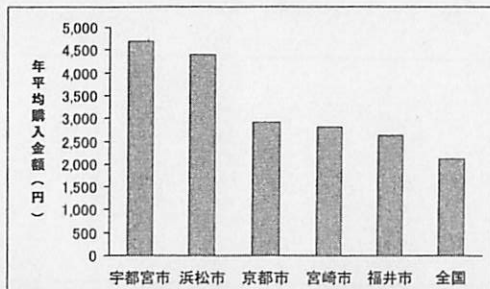
② $\frac{1}{\sqrt{6.4}\sqrt{5.2}} \left\{ \frac{(1-5)(3-6) + 2 \times (4-5)(7-6) + (7-5)(8-6) + (10-5)(9-6)}{20} \right\}$

③ $\frac{1}{6.4 \times 5.2} \left\{ \frac{(1-5)(3-6) + (4-5)(7-6) + (7-5)(8-6) + (10-5)(9-6)}{20} \right\}$

④ $\frac{1}{6.4 \times 5.2} \left\{ \frac{(1-5)(3-6) + 2 \times (4-5)(7-6) + (7-5)(8-6) + (10-5)(9-6)}{20} \right\}$

⑤ $\frac{1}{6.4 \times 5.2} \{ (1-5)(3-6) + 2 \times (4-5)(7-6) + (7-5)(8-6) + (10-5)(9-6) \}$

問15 新聞で見た「ぎょうざ日本一の街」について、総務省「家計調査」からさらに調べた。次のグラフは、2人以上の世帯が2009～2011年にぎょうざを購入した金額の年平均値について、都道府県庁所在市および政令指定都市のうち、購入金額の多い上位5市と全国を示したグラフである。



資料：総務省「家計調査」

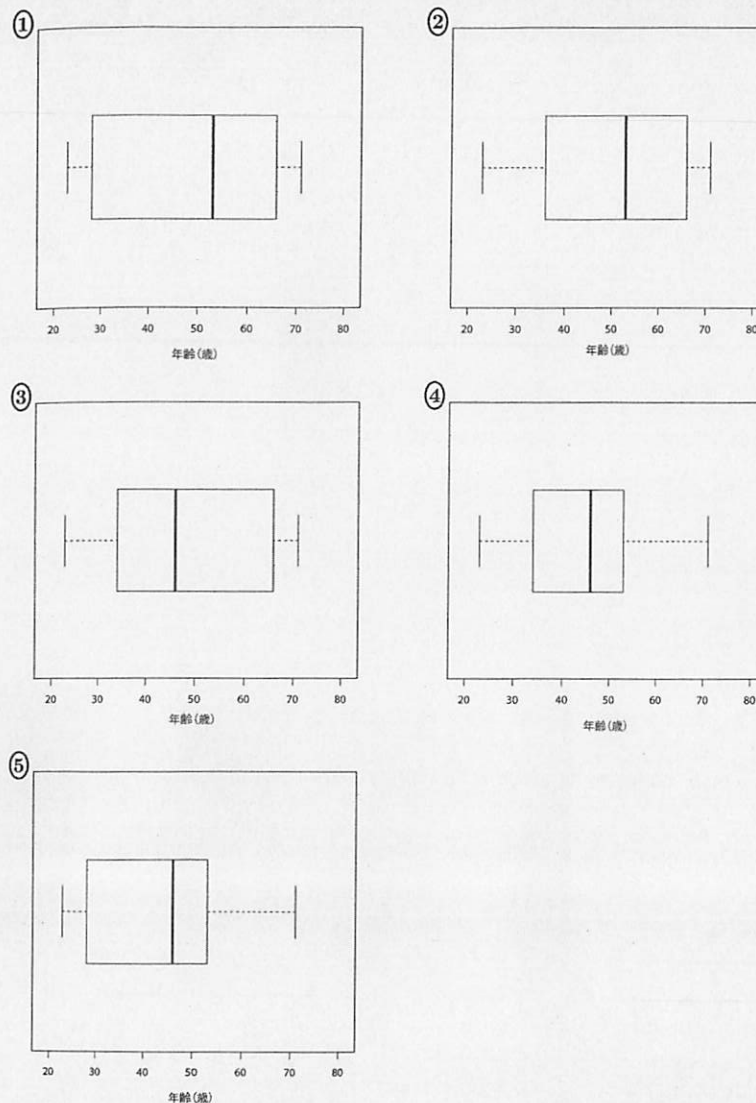
この結果から読み取れることとして正しいものを、次の①～⑤のうちから一つ選べ。 18

- ① 都道府県庁所在市および政令指定都市で全国よりもぎょうざの年平均購入金額が多いのは宇都宮市、浜松市、京都市、宮崎市、福井市の5つの市のみである。
- ② ぎょうざの購入金額の2人以上の世帯の総計について、宇都宮市は京都市の1.5倍以上である。
- ③ 宇都宮市も浜松市も2人以上の世帯の半分以上がぎょうざを年間4,000円以上購入している。
- ④ 都道府県庁所在市および政令指定都市の中で、2009～2011年における世帯のぎょうざの平均購入金額が最も多いのは宇都宮市であり、全国平均の2倍以上である。
- ⑤ 宇都宮市の2人以上の世帯はどの世帯も、全国平均よりも多くぎょうざを購入している。

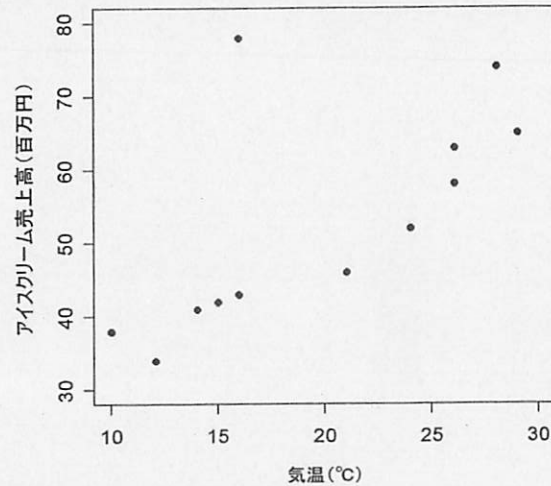
問16 次の表は、あるスポーツ施設に4月に新しく入会した会員40人の年齢を階級ごとにまとめたものである。

年齢	度数	累積度数
20歳以上～30歳未満	12	12
30歳以上～40歳未満	3	15
40歳以上～50歳未満	3	18
50歳以上～60歳未満	10	28
60歳以上～70歳未満	10	38
70歳以上～80歳未満	2	40

表を作成するもとなった40人のデータに対応する年齢の箱ひげ図として最も適切なものを、次の①～⑤のうちから一つ選べ。 19



- 問18 アイスクリームの売上高は、気温による影響があり気温が高い月ほど売上高が多いという仮説を立てた。このことを調べるために、ある地区のチェーン店における1月から12月までの月別の平均気温(°C)とアイスクリームの売上高(百万円)のデータの散布図を作成した。



上の散布図のデータに従えば、月別の平均気温の平均値は19.75°C、月別のアイスクリームの売上高の平均値は52.8百万円であった。また相関係数は0.676であり、予想よりも小さかった。ところが、観測値の入力に間違いがあることがわかり、図中の平均気温が16°C、アイスクリームの売上高が78百万円のデータに対して、正しい売上高は48百万円であった。

- [1] 正しい数値に直したとき、平均気温の平均値とアイスクリームの売上高の分散はどう変化するか、次の①～⑤のうちから一つ選べ。 22

- ① 修正前よりも平均気温の平均値は大きくなり、売上高の分散は大きくなる。
- ② 修正前よりも平均気温の平均値は大きくなり、売上高の分散は小さくなる。
- ③ 修正前よりも平均気温の平均値は変わらず、売上高の分散は大きくなる。
- ④ 修正前よりも平均気温の平均値は変わらず、売上高の分散は小さくなる。
- ⑤ 修正前よりも平均気温の平均値は変わらず、売上高の分散は変わらない。

- [2] 正しい数値に直したとき、相関係数はどう変化するか、次の①～⑤のうちから一つ選べ。 23

- ① 修正前の相関係数よりも小さくなる。
- ② 修正前の相関係数と変わらない。
- ③ 修正前の相関係数の絶対値は変わらないが、符号は逆になる。
- ④ 修正前の相関係数よりも大きくなる。
- ⑤ これらの情報だけでは判断できない。

- [3] 正しい数値に直したとき、この結果について述べた次の記述のうち、最も適切なものを、次の①～⑤のうちから一つ選べ。 24

- ① 平均気温とアイスクリームの売上高の散布図から、観測値は直線状に分布しておらず、相関係数の絶対値もあまり高くないことから、気温が高い月ほど売上高が多い傾向があるとは言えない。
- ② 平均気温とアイスクリームの売上高の散布図から、観測値は右上がりの直線状に分布しており、相関係数からも正の相関がみられ、気温が高い月ほど売上高が多い傾向があった。
- ③ 平均気温とアイスクリームの売上高の散布図から、観測値は右下がりの直線状に分布しており、相関係数からも負の相関がみられ、気温が高い月ほど売上高が多い傾向があった。
- ④ 平均気温とアイスクリームの売上高の散布図から、いくつかの観測値が右上がりの直線状からはずれて分布しているため、気温が高い月ほど売上高が多い傾向があるとは言えない。
- ⑤ 平均気温とアイスクリームの売上高の散布図から、いくつかの観測値が右下がりの直線状からはずれて分布しているため、気温が高い月ほど売上高が多い傾向があるとは言えない。