〔2〕この箱ひげ図から読み取れることとして不適切なものを，次の①～⑤のうちから一つ選べ。 **2**

① 600 トン以上の都道府県の数は，東日本と西日本で同数である。

② 100 トン以下の都道府県の数は，西日本の方が多い。

③ 範囲は西日本の方が大きい。

④ 第 1 四分位数は東日本の方が大きい。

⑤ 西日本の府県の中に平均を 100 トン以上押し上げる外れ値があるので，平均は西日本の方が大きいと推察できる。

問 2 国際通貨基金（IMF）によると，南米のベネズエラでは，2010 年の物価は 2009 年と比べて 1.282 倍，すなわち 2010 年の 1 年当たり物価上昇率は 28.2 %であった。同様に，2011 年 ～ 2013 年の 1 年当たり物価上昇率がそれぞれ 26.1 %，21.1 %，40.6 %であった。2010 年 ～ 2013 年の 4 年間の平均物価上昇率は何%か。次の①～⑤のうちから最も適切なものを一つ選べ。 **3**

① 28.2 %　② 28.8 %　③ 29.0 %　④ 29.2 %　⑤ 29.4 %

問 3 あるスーパーマーケットで平日に買い物に来た客の 1 回の購入金額を食料品フロアと衣料品フロアで調べてみた。次の表は購入金額（税抜）に関して要約したものである。

	食料品（円）	衣料品（円）
最小値	10	300
第 1 四分位数	700	2,400
中央値	1,400	3,800
第 3 四分位数	2,500	8,000
最大値	30,000	120,000
平均	1,900	5,200
標準偏差	2,500	7,200
客数（人）	11,000	2,500

〔1〕購入金額の分布は食料品でも衣料品でもほぼ同じ形状をしていた。分布の形状に関して，次の①～⑤のうちから最も適切なものを一つ選べ。 **4**

① 第 1 四分位数と第 3 四分位数の間に観測値のほとんどが含まれる，一様型に近い形をしている。

② 中央値を中心とした左右対称のベル型をしている。

③ 平均を中心とした左右対称のベル型をしている。

④ 右の裾が長い形状をしている。

⑤ 左の裾が長い形状をしている。

〔2〕第 1 四分位数と第 3 四分位数を用いたばらつきの説明に関して，次の①～⑤のうちから最も適切なものを一つ選べ。 **5**

① 第 1 四分位数と第 3 四分位数の間には観測値のうち約四分の一が含まれる。

② 第 1 四分位数と第 3 四分位数の間には観測値のうち約半分が含まれる。

③ 第 1 四分位数と第 3 四分位数の間には観測値のうち約四分の三が含まれる。

④ 第 1 四分位数と第 3 四分位数の間には観測値のうちほとんどが含まれる。

⑤ 第 1 四分位数より小さい観測値は全体の四分の一，第 3 四分位数より大きい観測値は四分の三だが，その間に入る観測値の割合はわからない。

〔3〕次の文章は食料品と衣料品での購入金額のばらつきについて述べたものである。（ア）～（オ）にあてはまるものとして，下の①～⑤のうちから最も適切なものを一つ選べ。 **6**

購入金額の標準偏差は食料品は 2,500 円，衣料品は 7,200 円だから，購入金額のばらつきは衣料品の方が大きいということができる。しかし，食料品に比べて衣料品は単価の高いものが多く，平均も食料品と比べて約 2.7 倍であるから，衣料品の標準偏差が大きくなることは当然である。このような場合に観測値のばらつきを比較するには（ア）を用いればよい。（ア）の定義式は（イ）である。食料品の（ア）は（ウ），衣料品の（ア）は（エ）だから，購入金額のばらつきが相対的に大きいのは（オ）であるが，大きな差ではない。

① （ア）変動係数　（イ）$\dfrac{\text{平均}}{\text{標準偏差}}$　（ウ）0.76　（エ）0.72　（オ）衣料品

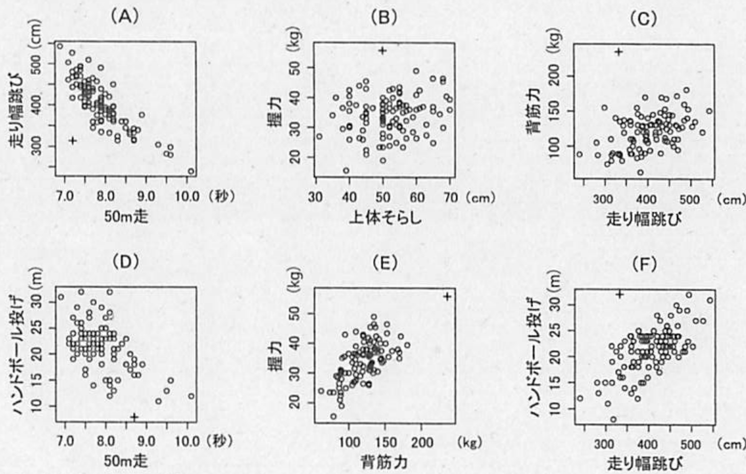
② （ア）変動係数　（イ）$\dfrac{\text{標準偏差}}{\text{平均}}$　（ウ）1.32　（エ）1.38　（オ）衣料品

③ （ア）変動係数　（イ）$\dfrac{\text{標準偏差}}{\text{平均}}$　（ウ）1.32　（エ）1.38　（オ）食料品

④ （ア）標準得点　（イ）$\dfrac{\text{平均}}{\text{標準偏差}}$　（ウ）0.76　（エ）0.72　（オ）食料品

⑤ （ア）標準得点　（イ）$\dfrac{\text{標準偏差}}{\text{平均}}$　（ウ）1.32　（エ）1.38　（オ）衣料品

問4 次の 6 つの散布図は，中学生 104 人に対して計測した運動能力のデータから作成したものである。なお，各散布図上に一つずつ記された「+」マークも観測値の一つであり，各散布図上で集団から比較的離れている点を表しているが，すべて同一の生徒というわけではない。



〔1〕各図の相関係数は

（ア）−0.807 （イ）−0.553 （ウ）0.197 （エ）0.360 （オ）0.537 （カ）0.657

のいずれかである。図と相関係数の正しい組合せはどれか。次の ① 〜 ⑤ のうちから最も適切なものを一つ選べ。 **7**

① A：ア　B：ウ　C：エ　D：イ　E：カ　F：オ
② A：カ　B：ウ　C：オ　D：エ　E：ア　F：イ
③ A：ア　B：エ　C：ウ　D：イ　E：オ　F：カ
④ A：ア　B：イ　C：ウ　D：エ　E：カ　F：オ
⑤ A：イ　B：カ　C：ウ　D：ア　E：エ　F：オ

〔2〕「+」の観測値を除いたときに相関係数の絶対値が最も増加する図と，最も減少する図はそれぞれどれか。次の ① 〜 ⑤ のうちから最も適切なものを一つ選べ。

8

① 最も増加する：D　最も減少する：A
② 最も増加する：B　最も減少する：A
③ 最も増加する：E　最も減少する：C
④ 最も増加する：F　最も減少する：B
⑤ 最も増加する：C　最も減少する：E

問5 次の表は，あるカレーライス専門店における客 400 人（男性 200 人，女性 200 人）を対象に実施したアンケートをまとめたクロス集計表である。この店では辛口と甘口の 2 種類のカレーライスを扱っている。

表 1：全顧客

	満足	不満	合計
辛口を注文	116	84	200
甘口を注文	124	76	200
合計	240	160	400

表 2：女性客のみ

	満足	不満	合計
辛口を注文	38	2	40
甘口を注文	122	38	160
合計	160	40	200

〔1〕男性客の中で辛口を注文して満足した客は何人いるか。次の ① 〜 ⑤ のうちから適切なものを一つ選べ。 **9**

① 2 人　② 38 人　③ 78 人　④ 82 人　⑤ 84 人

〔2〕この表から分かることの説明として，次の ① 〜 ⑤ のうちから最も適切なものを一つ選べ。 **10**

① 全顧客の集計では，辛口でも甘口でも満足の割合は 60 %程度であまり変わらない。したがって，性別ごとにわけて集計する必要は無く，性別に関わらず辛口と甘口には満足度の傾向にほとんど差がないことがわかる。

② 甘口を注文した客のうち，満足の割合は女性客では 75 %程度，男性客では 5 %程度である。満足の割合がともに 0 %でないということは，性別によらず辛口よりも甘口の方が好まれていることを意味している。

③ 女性客では，辛口の満足の割合は 95 %だが，甘口は 75 %程度である。男性客では，辛口の満足の割合は 50 %程度，甘口は 5 %である。性別によらず辛口の満足度が高いため，甘口よりも辛口を注文した人の方が満足度が高いといえる。

④ 女性で満足した人のうち甘口を注文したのは 75 %程度なのに対し，男性で満足した人の中で甘口を注文した客はほとんどいなかった。したがって，甘口を注文する割合は性別によって大きく異なることがわかる。

⑤ 満足した客の割合は，女性客では 80 %であったが男性客では 40 %にとどまった。したがって，辛口と甘口にわけて集計する必要は無く，男性の方が女性よりも不満であったことがわかる。

問 14 次の文章は，母比率 p の信頼区間について述べたものである。

標本の大きさを n，標本比率を $\hat{p}$ とする。$\hat{p}$ は確率変数であり，n が十分大きいとき平均 p，標準偏差 $\sqrt{p(1-p)/n}$ の正規分布にほぼ従う。したがって，$\hat{p}$ を標準化した確率変数 $Z=$（ア）は標準正規分布にほぼ従うので，$-1.96 \leq$（ア）≤ 1.96 が 95 % の確率で成り立つ。これを変形すると $\hat{p}-$（イ）$\leq p \leq \hat{p}+$（イ）となり，この区間が p を含む確率は 95% であることがわかる。この（イ）には未知の値 p が含まれるため，p の代わりに標本比率 $\hat{p}$ を用いることで p の近似的な信頼区間が得られる。

一方，標本比率を用いなくても，信頼区間のおおよその幅を見積もることができる。$p(1-p)$ が最大となるのは $p=$（ウ）のときであり，その最大値は（エ）である。そして，1.96 をほぼ 2 とみなすことにより，（イ）の上限はほぼ（オ）となることがわかる。したがって，信頼区間の幅が $2\times$（オ）以下であることがわかる。

〔1〕文中の（ア），（イ）にあてはまるものとして，次の ①～⑤ のうちから最も適切なものを一つ選べ。 **25**

① （ア）$\dfrac{\hat{p}-p}{\sqrt{p(1-p)}}$　（イ）$1.96\sqrt{np(1-p)}$

② （ア）$\dfrac{\hat{p}-p}{\sqrt{p(1-p)/n}}$　（イ）$1.96\sqrt{p(1-p)}$

③ （ア）$\dfrac{\hat{p}-p}{\sqrt{p(1-p)/n}}$　（イ）$1.96\sqrt{p(1-p)/n}$

④ （ア）$\dfrac{\hat{p}-p}{\sqrt{np(1-p)}}$　（イ）$1.96\sqrt{np(1-p)}$

⑤ （ア）$\dfrac{\hat{p}-p}{\sqrt{np(1-p)}}$　（イ）$1.96\sqrt{p(1-p)/n}$

〔2〕文中の（ウ）～（オ）にあてはまるものとして，次の ①～⑤ のうちから最も適切なものを一つ選べ。 **26**

① （ウ）1　（エ）0　（オ）$\sqrt{n}$

② （ウ）1　（エ）$\dfrac{1}{4}$　（オ）$\dfrac{1}{\sqrt{n}}$

③ （ウ）0.5　（エ）$\dfrac{1}{4}$　（オ）$\sqrt{n}$

④ （ウ）0.5　（エ）$\dfrac{1}{4}$　（オ）$\dfrac{1}{\sqrt{n}}$

⑤ （ウ）0　（エ）0　（オ）1

問 15 ある刺激を与えたときの血圧（収縮期血圧）の変化を調べるために，10 人の被験者に対して，刺激を与える前の血圧（mmHg）と刺激を与えた後の血圧（mmHg）を測定した。

No.	1	2	3	4	5	6	7	8	9	10	平均
刺激前	130	118	128	135	126	120	126	140	127	130	128.0
刺激後	135	120	132	135	129	128	135	139	135	132	132.0

〔1〕この刺激を与えた後に血圧が上がる変化があるかを有意水準 5 % で片側検定したい。用いる t 分布の自由度と棄却域について，次の ①～⑤ のうちから最も適切なものを一つ選べ。 **27**

① 自由度 9　棄却域 $t \geq 1.833$

② 自由度 9　棄却域 $|t| \geq 2.262$

③ 自由度 10　棄却域 $|t| \geq 2.228$

④ 自由度 18　棄却域 $t \geq 1.734$

⑤ 自由度 18　棄却域 $|t| \geq 2.101$

〔2〕t 統計量の値と検定の結果について，次の ①～⑤ のうちから最も適切なものを一つ選べ。 **28**

① t 統計量の値は 1.51 であり，帰無仮説は棄却されず，この刺激を与えた後に血圧が上がる変化があるとはいえない，と判断する。

② t 統計量の値は 1.51 であり，帰無仮説は棄却されず，この刺激を与えた後に血圧が上がる変化がないとはいえない，と判断する。

③ t 統計量の値は 3.65 であり，帰無仮説を棄却し，この刺激を与えた後に血圧が上がる変化がないとはいえない，と判断する。

④ t 統計量の値は 3.65 であり，帰無仮説を棄却し，この刺激を与えた後に血圧が上がる変化があるとはいえない，と判断する。

⑤ t 統計量の値は 3.65 であり，帰無仮説を棄却し，この刺激を与えた後に血圧が上がる変化がある，と判断する。

問8　100円玉5枚，10円玉7枚，1円玉3枚の入った小銭入れから，同時に3枚の硬貨を取り出す。いずれの硬貨を取り出すのも同様に確からしいとする。

〔1〕取り出した3枚の金額の合計が111円である確率はいくらか。次の①～⑤のうちから適切なものを一つ選べ。 16

① $\frac{1}{3}$　② $\frac{1}{26}$　③ $\frac{1}{105}$　④ $\frac{1}{5}$　⑤ $\frac{3}{13}$

〔2〕取り出した3枚の金額の合計が150円以上である確率はいくらか。次の①～⑤のうちから適切なものを一つ選べ。 17

① $\frac{1}{3}$　② $\frac{1}{25}$　③ $\frac{15}{91}$　④ $\frac{22}{91}$　⑤ $\frac{20}{273}$

〔3〕取り出した3枚の金額の合計が150円以上であるという条件のもとで，その3枚の中に1円玉が含まれる条件付き確率はいくらか。次の①～⑤のうちから適切なものを一つ選べ。 18

① $\frac{1}{3}$　② $\frac{3}{11}$　③ $\frac{8}{11}$　④ $\frac{6}{91}$　⑤ $\frac{8}{91}$

問9　4つの値をとる離散型確率変数 X の確率関数 $f(x)$ が

$$f(1)=0.2,\quad f(2)=0.4,\quad f(5)=0.25,\quad f(10)=0.15$$

であるとき，期待値 $E(X)$ はいくらか。次の①～⑤のうちから適切なものを一つ選べ。 19

① 0.9375　② 2.35　③ 3.25　④ 3.75　⑤ 4.5

問10　ある都市におけるPM2.5（微小粒子状物質）の1日平均量 X は独立に正規分布に従い，その平均は $20\,\mu g/m^3$，標準偏差は $5\,\mu g/m^3$ であると仮定する。世界保健機関（WHO）のガイドラインでは，PM2.5の1日平均量に対する指針値を $25\,\mu g/m^3$ と定めている。

〔1〕PM2.5の1日平均量 X が指針値 $25\,\mu g/m^3$ を超える確率はいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。 20

① 2.3 %　② 16 %　③ 32 %　④ 46 %　⑤ 84 %

〔2〕1日平均量 X の1週間（7日）での平均が指針値 $25\,\mu g/m^3$ を超える確率はいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。 21

① 0.1 %　② 0.4 %　③ 0.7 %　④ 1.5 %　⑤ 2.3 %

〔3〕1週間のうち，1日平均量 X が指針値 $25\,\mu g/m^3$ を超える日が2日以上となる確率はいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。 22

① 1.0 %　② 29 %　③ 31 %　④ 69 %　⑤ 71 %

問11　二者択一で答えさせる36問の問題があり，24問以上の正解で合格である。このとき，全くでたらめの解答をしたときの合格率はいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。 23

① 50 %　② 30 %　③ 10 %　④ 5 %　⑤ 3 %

問 12 A さんは日常生活の風景をデジタルカメラで写真におさめ，パソコンに保存するという趣味を持っている。A さんが先週 1 週間に撮影した写真 31 枚について，それらのファイルサイズ（単位はメガバイト）の標本平均および不偏分散を計算したところ，次のようになった。

$$\text{標本平均} = 3.24, \quad \text{不偏分散} = 0.04$$

1 枚の写真のファイルサイズは独立に平均 μ，未知の分散 σ^2 の正規分布に従うと仮定して，母平均 μ の 95 %信頼区間を求める式はどれか。次の ①～⑤ のうちから適切なものを一つ選べ。 $\boxed{24}$

① $\left[3.24 - 1.697\sqrt{0.04/30},\ 3.24 + 1.697\sqrt{0.04/30}\right]$

② $\left[3.24 - 1.96\sqrt{0.04/31},\ 3.24 + 1.96\sqrt{0.04/31}\right]$

③ $\left[3.24 - 1.96\sqrt{0.04/30},\ 3.24 + 1.96\sqrt{0.04/30}\right]$

④ $\left[3.24 - 2.042\sqrt{0.04/31},\ 3.24 + 2.042\sqrt{0.04/31}\right]$

⑤ $\left[3.24 - 2.042\sqrt{0.04^2/30},\ 3.24 + 2.042\sqrt{0.04^2/30}\right]$

問 13 次の文章は母平均の検定について述べたものである。

母集団分布は平均 μ，分散 σ^2 の正規分布であり，σ^2 は未知であるとする。大きさ n の無作為標本に基づき，帰無仮説 $H_0 : \mu = \mu_0$ を有意水準 1 %で検定したい。対立仮説は $H_1 : \mu \neq \mu_0$ とする。通常は（ア）検定を用いる。この方法では（ア）統計量の絶対値が（イ）分布の上側 0.5 %点以上であれば，帰無仮説は（ウ）。

一方，データとは無関係に一様乱数（一様分布 $U(0,1)$ に従う確率変数の実現値）を求め，その値が 0.01 以下であれば帰無仮説 H_0 を棄却するという方法も有意水準 1 %の検定である。この方法が（ア）検定と比べて不合理なのは，（エ）を全く考慮していないことである。一様乱数を用いた検定では，どのような対立仮説の下でも，（エ）をおかす確率は（オ）である。（ア）検定の場合は，この確率は常に（オ）以下である。

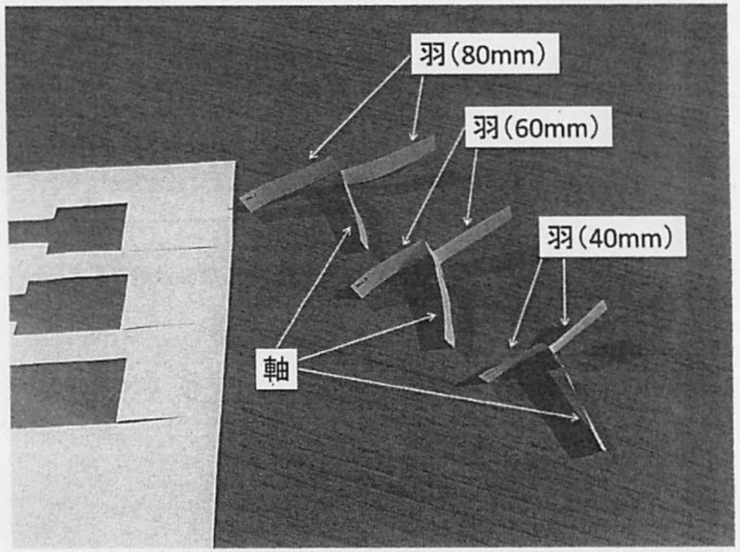
〔1〕文中の（ア）～（ウ）にあてはまるものの組合せとして，次の ①～⑤ のうちから適切なものを一つ選べ。 $\boxed{25}$

① （ア）t　（イ）自由度 n の t　（ウ）棄却されない
② （ア）t　（イ）自由度 $n-1$ の t　（ウ）棄却されない
③ （ア）t　（イ）自由度 $n-1$ の t　（ウ）棄却される
④ （ア）z　（イ）標準正規　（ウ）棄却されない
⑤ （ア）z　（イ）標準正規　（ウ）棄却される

〔2〕文中の（エ），（オ）にあてはまるものの組合せとして，次の ①～⑤ のうちから適切なものを一つ選べ。 $\boxed{26}$

① （エ）第一種の過誤　（オ）0.99
② （エ）第一種の過誤　（オ）0.01
③ （エ）第二種の過誤　（オ）0.99
④ （エ）第二種の過誤　（オ）0.01
⑤ （エ）第二種の過誤　（オ）0.50

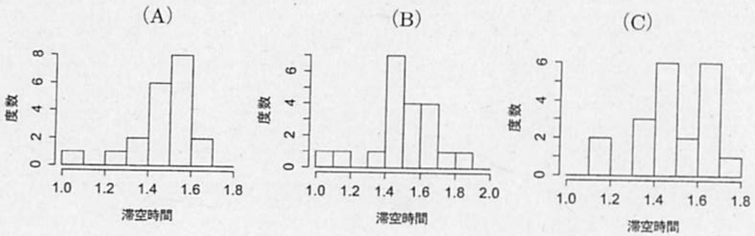
問 2　A 君は，紙コプター（次図）を作成し，滞空時間について調べる実験を行った。紙コプターは，宙に放すと軸を中心に羽が回りながらゆっくり落下する手作りおもちゃであり，別名，紙ヘリコプターとも呼ばれるものである。

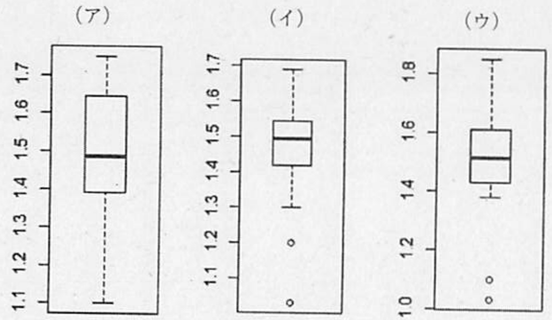


その実験では，翼長のみ変え，40mm, 60mm, 80mm としたものについてそれぞれ 20 回ずつの滞空時間（秒）を測定した。翼長 40mm について滞空時間の測定値は以下の通りである。

1.30,　1.48,　1.50,　1.48,　1.45,　1.44,　1.53,　1.55,　1.49,　1.40
1.57,　1.52,　1.52,　1.65,　1.58,　1.20,　1.69,　1.54,　1.38,　1.03

翼長 40mm, 60mm, 80mm の滞空時間についてヒストグラムを作成したものが（A），（B），（C）のいずれかである。ただし，ヒストグラムの各階級は 1.0 秒以上 1.1 秒未満のように 0.1 秒間隔で下限値は含み，上限値は含まないものとする。また，この 3 つのデータについて箱ひげ図を作成したものが（ア），（イ），（ウ）のいずれかである。この箱ひげ図では，"「第 1 四分位数」―「四分位範囲」× 1.5" 以上の値を取る測定値の最小値までひげを引き，それより小さな値を外れ値として ○ で示している。





〔1〕翼長 40mm のデータに対するヒストグラムと箱ひげ図の組合せはどれか。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **2**

① （A）－（ア）　② （A）－（イ）　③ （B）－（ア）
④ （B）－（ウ）　⑤ （C）－（ア）

〔2〕この翼長 40mm のデータについて中央値と平均値を求めた場合に，その大小関係はどのようになるか。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。
3

① 平均値 ＝ 中央値　② 平均値 ＞ 中央値
③ 平均値 ＜ 中央値　④ 中央値 ＝ 2 × 平均値
⑤ 2 つの峰があると考えられるので平均値は求められない。

2014年11月 問題

問 4　ある大都市で，市民スポーツセンターの利用者が多いことからスポーツセンターの大型化を計画している。そのため利用者数の予想を立てる必要が生じた。そこで利用者調査として，まず市内の区の人口に応じて区の確率比例抽出を行い，抽出された区について 10 世帯につき 1 世帯を調査する方法をとった。この調査法の枠組みとして最も適切なものを次の ①～⑤ のうちから一つ選べ。 **5**

① 単純無作為抽出法　② 集落抽出法　③ 層別抽出法
④ 多段抽出法　⑤ 全数調査

問 5　ある会社の待合室で，男性 8 人，女性 7 人の社員が会議のために部長を待ち，コーヒーか紅茶のどちらかを飲んでいた。男性でコーヒーを飲んでいたのは 5 人，女性でコーヒーを飲んでいたのは 2 人であった。

〔1〕部長が男性の 1 人にコーヒーか紅茶のどちらを飲んでいたか聞いたとき，その人が飲んでいたものがコーヒーであった確率はいくらか。次の ①～⑤ のうちから適切なものを一つ選べ。 **6**

① $\frac{3}{8}$　② $\frac{5}{8}$　③ $\frac{2}{7}$　④ $\frac{5}{7}$　⑤ $\frac{3}{15}$

〔2〕部長が男性 1 人，女性 1 人に飲んでいたものを聞いたときに，2 人が違う飲み物を飲んでいた確率はいくらか。次の ①～⑤ のうちから適切なものを一つ選べ。 **7**

① $\frac{3}{7}$　② $\frac{8}{15}$　③ $\frac{25}{56}$　④ $\frac{31}{56}$　⑤ $\frac{31}{105}$

問 6　次の表はある地域における 1 日の死亡者数の集計結果（500 日間）である。

死亡者数（人）	0	1	2	3	4	5	6 以上	計
件数（日数）	55	144	140	95	45	15	6	500

〔1〕死亡者数の平均値はいくらか。ただし，死亡者数が 7 人以上の日はなかった。次の ①～⑤ のうちから最も適切なものを一つ選べ。 **8**

① 1.50　② 1.75　③ 2.00　④ 2.25　⑤ 2.50

〔2〕1 日の死亡者数 X がパラメータ λ のポアソン分布にしたがうと仮定するとき，ある日の死亡者数が 3 人である確率を求める式として正しいものはどれか。次の ①～⑤ のうちから適切なものを一つ選べ。 **9**

① λ　② $\lambda(1-\lambda)^3$　③ ${}_6C_3\lambda^3(1-\lambda)^3$
④ $\lambda e^{-3\lambda}$　⑤ $\frac{\lambda^3 e^{-\lambda}}{3!}$

〔3〕1 日の死亡者数 X がパラメータ λ のポアソン分布にしたがうと仮定するとき，X^2 の期待値 $E(X^2)$ とパラメータ λ の関係として正しいものはどれか。次の ①～⑤ のうちから適切なものを一つ選べ。 **10**

① $E(X^2)=\frac{2}{\lambda^2}$　② $E(X^2)=\frac{1}{\lambda}$　③ $E(X^2)=\lambda$
④ $E(X^2)=\lambda^2+\lambda$　⑤ $E(X^2)=\lambda^2$

〔4〕1 日の死亡者数 X がパラメータ λ のポアソン分布にしたがっているか否かの検定を行う。データの平均値をもとにパラメータを推定し，次の表の期待度数を得た。

死亡者数（人）	0	1	2	3	4	5	6 以上
期待度数（日）	67.7	135.3	135.3	90.2	45.1	18.0	8.3

このとき適合度検定の判断として，次の ①～⑤ のうちから最も適切なものを一つ選べ。 **11**

① 検定統計量の χ^2 値 =0.647 が自由度 7 の χ^2 分布の上側 5%点より大きいか比べる。
② 検定統計量の χ^2 値 =0.647 が自由度 6 の χ^2 分布の上側 5%点より大きいか比べる。
③ 検定統計量の χ^2 値 =4.498 が自由度 7 の χ^2 分布の上側 5%点より大きいか比べる。
④ 検定統計量の χ^2 値 =4.498 が自由度 6 の χ^2 分布の上側 5%点より大きいか比べる。
⑤ 検定統計量の χ^2 値 =4.498 が自由度 5 の χ^2 分布の上側 5%点より大きいか比べる。

問 7 あるスポーツチームにおける選手の身長 (cm) は平均が 174，分散が 9^2 の正規分布にしたがう。この母集団から，16 人の選手を無作為に抽出する。

〔1〕抽出された 16 人のうち少なくとも 1 人が 184cm 以上である確率を求めたい。Z が標準正規分布にしたがう確率変数であるとするとき，その確率を求める式として正しいものはどれか。次の ①～⑤ のうちから適切なものを一つ選べ。 **12**

① $\left\{P\left(\dfrac{184-174}{9} \leq Z\right)\right\}^{16}$　② $1-\left\{P\left(Z<\dfrac{184-174}{9}\right)\right\}^{16}$

③ $P\left(\dfrac{184-174}{9} \leq Z\right) \times 16$　④ $1-P\left(Z<\dfrac{184-174}{9}\right) \times 16$

⑤ $\left\{1-P\left(Z<\dfrac{184-174}{9}\right)\right\}^{16}$

〔2〕抽出された 16 人の身長の標本平均を考えるとき，その期待値と分散はそれぞれいくらか。次の ①～⑤ のうちから正しい組合せを一つ選べ。 **13**

① 期待値 174，分散 9^2　② 期待値 174，分散 $\dfrac{9^2}{16}$

③ 期待値 174^{16}，分散 81^{16}　④ 期待値 2784，分散 36

⑤ 期待値 2784，分散 $\dfrac{9^2}{16}$

問 8 1 から 8 までの目が等確率で出る正八面体のサイコロがある。このサイコロを 640 回振ったとき，1 の目が 90 回以上出る確率はいくらか。次の ①～⑤ のうちから最も適切なものを一つ選べ。 **14**

① 0.09　② 0.12　③ 0.15　④ 0.18　⑤ 0.21

問 9 あるバーガーショップで販売されているフライドポテトの説明には，重量は平均 100g，標準偏差 4g と表記されている。ある店舗で「購入したフライドポテトの重量が 94g であった。この店のフライドポテトの量は少ないのではないか。」とクレームをつける客がいた。フライドポテトの重量が正規分布にしたがうとして，以下の各問に答えよ。

〔1〕平均 100g，標準偏差 4g の表記が正しいとして，無作為にフライドポテトの重量を 100 個調べたとき，94g 以下となるものは何個あると期待されるか。次の ①～⑤ のうちから最も適切なものを一つ選べ。 **15**

① 0　② 1.5　③ 2　④ 7　⑤ 94

〔2〕この店舗でフライドポテトの重量を無作為に 100 個調べたところ，平均値は 98.8g であった。平均 100g，標準偏差 4g の表記が正しいとき，無作為に 100 個選んで重量を調べたとき，その標本平均が 98.8g 以下となる確率はいくらか。次の ①～⑤ のうちから最も適切なものを一つ選べ。 **16**

① 0.3821　② 0.2266　③ 0.0668　④ 0.0250　⑤ 0.0013

〔3〕〔2〕の調査結果に対して，どのように判断すればよいか。次の ①～⑤ のうちから最も適切なものを一つ選べ。 **17**

① 平均 100g，標準偏差 4g の母集団から無作為に 100 個取り出した標本平均が 98.8g となることは，誤差の範囲であるので，改善の必要はないと判断する。

② 平均 100g，標準偏差 4g の母集団から無作為に 100 個取り出した標本平均が 98.8g となることは，5%より大きな確率で起こる。したがって，このようなことは起こりえると考え，改善の必要はないと判断する。

③ 平均 100g，標準偏差 4g の母集団から無作為に 100 個取り出した標本平均が 98.8g となるのは，5%より小さな確率でしか起こらないことがわかった。したがって，偶然ではなく，何らかの原因で内容量が少なくなっていると考え，改善の必要があると判断する。

④ 平均 100g，標準偏差 4g の母集団から無作為に 100 個取り出した標本平均が 98.8g となることは，確率的にめったに起こらないことなので，改善の必要はないと判断する。

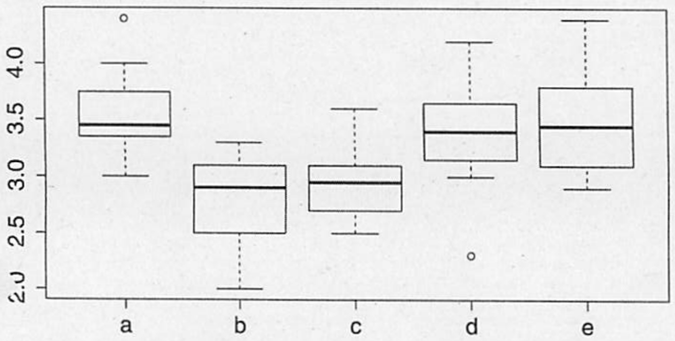
⑤ さまざまな場合が考えられ，今回の調査結果も確率を計算するまでもなく，可能性としてはありえる。そのため，改善の必要はないと判断する。

問 10　同じ植物の 2 つの品種 A と B がある。品種 A は 21 本，品種 B は 16 本についてある部位の長さ (cm) を測定した。下の表は，その平均値と分散である。

	平均値	分散
品種 A	2.78	0.145
品種 B	2.93	0.095

これより，2 つの品種のその部位の長さに差があるかどうかを検定したい。その部位の長さは正規分布にしたがい，分散は等しいと仮定できるものとして t 検定を行う。

〔1〕検定の前に，2 つの標本の分布の様子を箱ひげ図で確認しておく。次の図のうち，品種 A と品種 B はそれぞれどの箱ひげ図にあたるか。下の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **18**



① 品種 A：a，品種 B：e　② 品種 A：b，品種 B：c　③ 品種 A：a，品種 B：d

④ 品種 A：b，品種 B：e　⑤ 品種 A：c，品種 B：d

〔2〕2 つの標本をプールした分散 s^2 はいくらか。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **19**

① $s^2 = 0.120$　② $s^2 = 0.124$　③ $s^2 = 0.130$　④ $s^2 = 0.262$　⑤ $s^2 = 0.872$

〔3〕t 統計量の計算式はどれか。次の ① ～ ⑤ のうちから適切なものを一つ選べ。

20

① $t = \dfrac{2.78 - 2.93}{s}$　② $t = \dfrac{2.78 - 2.93}{\frac{s}{37}}$　③ $t = \dfrac{2.78 - 2.93}{s\sqrt{\frac{1}{21} + \frac{1}{16}}}$

④ $t = \dfrac{2.78 - 2.93}{s\sqrt{\frac{1}{35}}}$　⑤ $t = \dfrac{2.78 - 2.93}{\sqrt{\frac{s^2}{37}}}$

〔4〕この検定の結果として次のような結論を導いた。空欄（ア）～（ウ）に入る言葉として正しい組合せはどれか。下の ① ～ ⑤ のうちから適切なものを一つ選べ。 **21**

有意水準 5 %で両側検定を用いることとする。自由度 35 の t 分布の上側 2.5 %点は 2.030 である。〔3〕で求めた t 統計量の値は -1.284 であるので，「2 つの品種のある部位の長さに（ア）」という帰無仮説は（イ）。よって「2 つの品種のある部位の長さに（ウ）」と結論する。

① （ア）差はない　（イ）棄却される　（ウ）差があるといえる
② （ア）差はある　（イ）棄却される　（ウ）差があるといえる
③ （ア）差はある　（イ）棄却されない　（ウ）差があるといえる
④ （ア）差はない　（イ）棄却されない　（ウ）差があるとはいえない
⑤ （ア）差はある　（イ）棄却される　（ウ）差があるとはいえない

問17　ある県における政党Aの支持率について調査を行い，支持率の信頼区間を求めたところ，信頼区間の幅がやや広かった。次回の調査のときには信頼区間の幅を約半分にしたい。次の①～⑤のうちから適切な方法を一つ選べ。 **21**

① 標本の大きさ n を固定して，信頼係数を 2 倍にする。
② 標本の大きさ n を固定して，信頼係数を 1/2 倍にする。
③ 信頼係数を固定して，標本の大きさ n を 2 倍にする。
④ 信頼係数を固定して，標本の大きさ n を 4 倍にする。
⑤ 信頼係数を固定して，標本の大きさ n を 1/2 倍にする。

問18　ある母集団の母平均 μ についての 95 %信頼区間は (48, 65) であった。次の記述 a から c がある。

a　標本の 95%が含まれている区間が，48 から 65 である。
b　標本平均が 48 から 65 の間にある確率は 0.95 である。
c　同じ大きさを持つランダムな標本を取り，それぞれについて 95 %信頼区間を求める手続きを多数回行うと，μ は，この手順でできる区間のうちの約 95 %の区間に含まれる。

この記述について，次の①～⑤のうちから最も適切なものを一つ選べ。 **22**

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 bとcのみ正しい

問19　A市では，道路に生ゴミが放置されているという苦情が市役所に持ち込まれた。そこで，生ゴミの回収回数を増やすかどうかを検討するために，世帯ごとの1年間のゴミ廃棄量を調査した。市役所では，公開されている従来の平均と標準偏差を用いて，「帰無仮説 H_0:ゴミ廃棄量の平均は変化していない」について正規分布にもとづく z 検定を行った。市役所側は，有意水準 $\alpha = 0.05$ の両側検定を行ったところ，帰無仮説は棄却できなかったと住民に説明した。しかし，住民側が同じデータについて同じ有意水準で片側検定を行ったところ，帰無仮説は棄却された。市役所と住民側が計算して求めた z 値を，次の①～⑤のうちから一つ選べ。ただし，計算結果は小数点第3位で四捨五入している。 **23**

① 1.30　② 1.50　③ 1.90
④ 1.98　⑤ 2.25

問20　ある法案に賛成か反対かについて住民に調査を行ない，男女別に集計して，次の2行3列のクロス集計表を得た。

	賛成	反対	どちらでもない	行計 n_i
男性	102	48	50	200
女性	72	54	24	150
列計 m_j	174	102	74	N =350

〔1〕 帰無仮説「H_0:法案への賛否と性別は独立である」のもとでクロス集計表の各セルの期待度数を計算したい。次の①～⑤のうちから正しい式を一つ選べ。ただし，第 i 行 (i=1,2)，第 j 列 (j=1,2,3) での期待度数を e_{ij}，観測度数を n_{ij}，第 i 行の観測度数の和を $n_i = \sum_{j=1}^{3} n_{ij}$，第 j 列の観測度数の和を $m_j = \sum_{i=1}^{2} n_{ij}$，総和を $N = \sum_{i=1}^{2}\sum_{j=1}^{3} n_{ij}$ と表記する。 **24**

① $e_{ij} = n_{ij} + n_i + m_j - N$
② $e_{ij} = n_i m_j / N$
③ $e_{ij} = (n_{ij} + n_i + m_j)/6$
④ $e_{ij} = (2 \times 3)\sqrt{n_i}\sqrt{m_j}/\sqrt{N}$
⑤ $e_{ij} = (n_i/3 + m_j/2)/2$

〔2〕 法案への賛否と性別は独立であるかを見るため，上のデータについてカイ二乗統計量を求めたところ，$\chi^2 = 7.674$ であった。有意水準を 5 %とした検定結果として次の①～⑤のうちから適切なものを一つ選べ。 **25**

① χ^2 分布表の自由度 6 の行の値から，法案への賛否と性別は独立ではないと判断できる。

② χ^2 分布表の自由度 6 の行の値から，法案への賛否と性別は独立であると判断できる。

③ χ^2 分布表の自由度 3 の行の値から，法案への賛否と性別は独立ではないと判断できる。

④ χ^2 分布表の自由度 3 の行の値から，法案への賛否と性別は独立であると判断できる。

⑤ χ^2 分布表の自由度 2 の行の値から，法案への賛否と性別は独立ではないと判断できる。

問 21 ある大学の教員は，自分の担当する授業中に 60 名の学生全員の 1 分間の脈拍値を測定した。そして，翌週の同じ授業時間に再度同じ 60 名の学生の 1 分間の脈拍値を測定した。2 回の授業での出席学生は同じであり，脈拍値は近似的に正規分布に従うとする。第 1 回目の計測で脈拍値が高かった 30 名を高脈拍値群とし（平均 $\bar{x}_H$），低かった 30 名を低脈拍値群として（平均 $\bar{x}_L$），翌週の第 2 回目の各群の脈拍値の平均を計算したところ，高脈拍値群での第 2 回目の脈拍値の平均 $\bar{y}_H$ は低脈拍値群の第 2 回目の脈拍値の平均 $\bar{y}_L$ よりも高かったものの，第 1 回目の計測における脈拍値の平均 $\bar{x}_H$ よりは低かった。すなわち，$\bar{y}_H > \bar{y}_L$ であるが $\bar{y}_H < \bar{x}_H$ であった。このとき，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 $\boxed{26}$

① 同じ学生に対して 2 回脈拍値を測っていることから，当然 2 回の計測値はほぼ同じ値になるはずであり，今回の計測で $\bar{y}_H < \bar{x}_H$ となったのは脈拍の計測法に問題があったことを示唆している。

② 同じ学生に対して 2 回脈拍値を測っていることから，今回は偶然変動により $\bar{y}_H < \bar{x}_H$ であったが，この種の計測を何回も繰り返せば，ほぼ $\bar{y}_H = \bar{x}_H$ となるはずである。

③ 高脈拍値群で $\bar{y}_H < \bar{x}_H$ となったのであるから，低脈拍値群についても当然 $\bar{y}_L < \bar{x}_L$ となっている可能性が高い。

④ 高脈拍値群で $\bar{y}_H < \bar{x}_H$ となったのであるが，低脈拍値群については逆に 1 回目の計測よりも平均が高くなり $\bar{y}_L > \bar{x}_L$ となっている可能性が高い。

⑤ 高脈拍値群で $\bar{y}_H < \bar{x}_H$ となったことは，少なくとも脈拍が高い群では多くの学生が 2 回目の計測までの 1 週間の間に脈拍値を下げるための方策を取ったためと考えられる。

問 22 次のデータはＡ社とＢ社から納品された試薬の純度のパーセンテージを 8 名の分析者が測定して得た結果である。Ａ社の母平均を μ_A，Ｂ社の母平均を μ_B として，2 社間で試薬の純度の平均にどの位の差があるかを検討することになった。

分析者	1	2	3	4	5	6	7	8	平均	標準偏差
Ａ社 y_A	5.7	4.7	7.5	6.0	6.8	5.6	6.0	8.7	6.375	1.25
Ｂ社 y_B	6.3	5.4	8.6	5.4	7.6	5.6	5.9	7.2	6.500	1.18

〔1〕 データの解析を依頼されたＳ社員は，2 社間の平均値の差 $\mu_A - \mu_B$ について 95 %信頼区間を求めることにした。Ｓ社員は，2 社のこのデータの標準偏差は同じくらいであることから，この 2 群のデータを用いて「プールした分散」を計算したところ $s^2 = 1.477$ となった。このとき，独立な 2 群の平均の差の 95 %信頼区間を求める式を次の ① ～ ⑤ のうちから一つ選べ。 $\boxed{27}$

① $-0.125 \pm 1.960 \times 1.22$
② $-0.125 \pm 2.120 \times 1.477$
③ $-0.125 \pm 1.960 \times 1.22/2$
④ $-0.125 \pm 2.145 \times 1.22/2$
⑤ $-0.125 \pm 2.120 \times 1.22/2$

〔2〕 Ｓ社員から検討結果の報告を受けた上司のＴ部長は，同じ分析者が 2 回測定していて，分析者間に違いがあるかもしれないので対応のあるデータに関する区間推定法を用いるべきではないかと助言した。Ｓ社員は，分析者毎の測定値の差を求めたところ以下のような結果を得た。

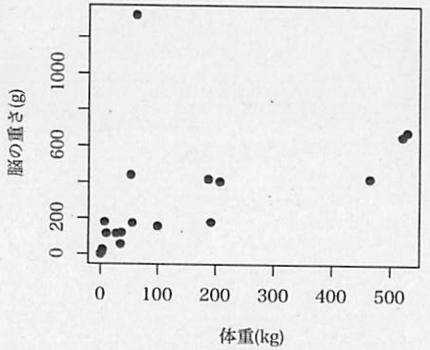
分析者	1	2	3	4	5	6	7	8	平均	標準偏差
差 d	-0.6	-0.7	-1.1	0.6	-0.8	0.0	0.1	1.5	-0.125	0.861

これらの値に基づく平均の差 $\mu_A - \mu_B$ の 95%信頼区間の計算式を次の ① ～ ⑤ のうちから一つ選べ。 $\boxed{28}$

① $-0.125 \pm 1.960 \times 0.861/\sqrt{8}$
② $-0.125 \pm 2.145 \times 0.861/\sqrt{8}$
③ $-0.125 \pm 2.365 \times 0.861/\sqrt{8}$
④ $\pm 1.960 \times 0.861/\sqrt{8}$
⑤ $\pm 2.145 \times 0.861/\sqrt{8}$

問 23 ある動物学者は，動物の体重 (kg) と脳の重さ (g) について，「体重が重くなると脳も重くなる」という仮説を立てた。そこで，象などの大型動物を除いた 23 種類の動物それぞれの体重の平均と脳の重さの平均について，散布図を作成し相関係数 r を求めたところ，$r = 0.54$ であった。

体重と脳の重さ



同僚の統計学教授に相談したところ，「自然対数 log を用いて，log(体重) を説明変数 x とし，log(脳の重さ) を応答変数（従属変数）y として回帰分析を行って検討しては」と助言された。そこで，この助言をもとに線形単回帰分析を適用した。

基本統計量	平均	標準偏差
log(体重)	2.56	2.94
log(脳の重さ)	4.07	2.32

ある統計解析ソフトによる最小 2 乗法による結果は次のとおりである。

残差

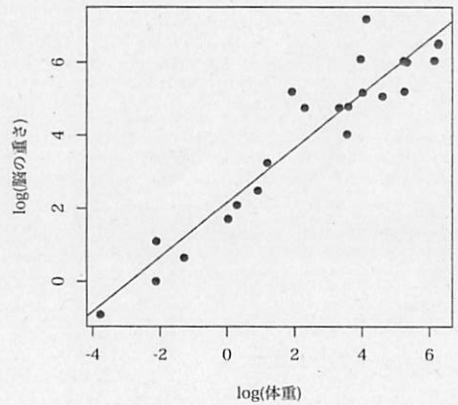
最小値	第 1 四分位数	中央値	第 3 四分位数	最大値
-0.88	-0.51	-0.26	0.15	1.95

回帰係数

	推定値	標準誤差	t 値	両側 p 値
切片	2.16	0.21	10.23	$< 1.29 \times 10^{-9}$
x	0.75	0.05	13.65	$< 6.52 \times 10^{-12}$

決定係数 R^2 0.90　自由度調整済み決定係数 0.89
F 値 186.4　自由度 (1,21)　p 値 6.52×10^{-12}

脳の重さについての回帰分析



〔1〕 log(体重の重さ) と log(脳の重さ) についての相関係数はいくらか。次の ① ～ ⑤ のうちから適切なものを一つ選べ。

① 0.54　② 0.75
③ 0.89　④ 0.90
⑤ 0.95

〔2〕 この結果からデータへの直線の当てはまりの程度を知りたい。どの値を用いるのが，最も適切か。次の ① ～ ⑤ のうちから一つ選べ。 $\boxed{29}$

① 切片の t 値　10.23　② 説明変数 x の回帰係数 0.75
③ 説明変数 x の t 値　13.65　④ 決定係数 R^2　0.90
⑤ F 値 186.4

〔3〕 この結果を用いると，log(体重) の値が大きい程，log(脳の重さ) が大きくなる，傾向が読み取れる。回帰式 $y = a + bx$ の係数を読み取り，体重 80 kg の動物の脳の重さを予測したい。脳の重さの予測値として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。ただし，log(80)=4.38 であり，e は自然対数の底とする。

$\boxed{30}$

① log(5.45)=1.70　② 5.45
③ 13.52　④ 62.16
⑤ $e^{5.45}$=231.60