この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

第7問 答案用紙<1>
(統 計 学)

問題 1

問 1 ラスパイレス価格指数

2017	2018	2019	2020	2021
100.00	108.66	109.65	103.58	110.16

問 2 パーシェ価格指数

2017	2018	2019	2020	2021
100.00	108.65	109.63	103.62	109.80

問 3

パーシェ価格指数は，ラスパイレス価格指数よりも低い値が得られる傾向にあるが，2018年度から2021年度までのエネルギー価格の動向は，どちらの価格指数でみても，2020年度には前年度よりも下落しているが，それ以外の年度では，前年度よりも上昇している。この結果，2021年度には，2017年度と比較して，エネルギー価格は10%程度上昇している。

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

第7問 答案用紙＜2＞
(統 計 学)

問題 2

問 1

平均値

平均値には，外れ値の影響を受けやすいという性質があるが，全データから求められるため，異なる年の貯蓄現在高のヒストグラムを比較するときには，平均値を用いるべきである。

中央値

中央値には，外れ値の影響を受けにくいという性質があるため，右に歪んだ貯蓄現在高のヒストグラムにおける中央に位置する値を知りたいときには，中央値を用いるべきである。

最頻値

貯蓄現在高のヒストグラムにおいて，平均値や中央値は，最頻値よりも大きな値となっており，多くの世帯の実感から離れたものとなっていると思われる。最頻値は，世帯の実感を考慮するために利用できる。

問 2

ヒストグラムでは，相対度数は，各棒の面積に比例するようにあらわされる。このため，貯蓄現在高のヒストグラムにおいて，標準級間隔100万円を棒の幅としたときに，棒の高さが相対度数を示すように縦軸の目盛りを設定していれば，棒の幅が標準級間隔より広い貯蓄現在高1,000万円以上の各階級については，相対度数を示す棒の高さは，縦軸の目盛りよりも低くなる。

問 3

3.96 %

問 4

1396 万円

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

第7問 答案用紙<3>
(統 計 学)

問題 3

問 1

$$\frac{1}{12}$$

問 2

$$\frac{1}{8}$$

問 3

$$\frac{1}{3}$$

問 4

$Y=b_2$のとき，参加者が最初に選択した箱に景品が入っている条件付確率は，つぎのように求められる。

$$\Pr\left(X=a \middle| Y=b_2\right)=\frac{\Pr\left(Y=b_2 \middle| X=a\right)\Pr\left(X=a\right)}{\Pr\left(Y=b_2\right)}=\frac{\frac{1}{3}\times\frac{1}{4}}{\frac{1}{3}}=\frac{1}{4}$$

一方，$Y=b_2$のとき，参加者がb_1に選択を変更することで景品を受け取る条件付確率は，つぎのように求められる。

$$\Pr\left(X=b_1 \middle| Y=b_2\right)=\frac{\Pr\left(Y=b_2 \middle| X=b_1\right)\Pr\left(X=b_1\right)}{\Pr\left(Y=b_2\right)}=\frac{\frac{1}{2}\times\frac{1}{4}}{\frac{1}{3}}=\frac{3}{8}$$

これらより，参加者は最初の選択を変更した方がよい。

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

【解答への道】

Ⅰ　合格ライン

問題 1

価格指数（ラスパイレス価格指数・パーシェ価格指数）に関する計算・説明問題である。ＴＡＣの答練でも繰り返し出題している内容であり，基本的な問題である。素点で７割程度の正答率が望まれる。

問題 2

ヒストグラムと代表値（平均値・中央値・最頻値）に関する問題である。標準的な問題であるが，記述問題に慣れていない方は前半部分がやや解きにくかったかもしれない。素点で６割程度の正答率が望まれる。

問題 3

一般に，モンティ・ホール問題とよばれる，確率に関する標準的な問題である。初見で感覚的に解こうとすると難しい部分もあったかもしれない。素点で６割程度の正答率が望まれる。

とくに，**問題 1** については，ＴＡＣの答練でも類題を繰り返し出題しており，得点すべき箇所での確実な得点が望まれる。全体として，第７問の合格ラインは，素点で６～７割程度と考えられる。

Ⅱ　答練との対応関係

問題 1

論文基礎答練第１回　第１問 **問題 1**　　論文応用答練第１回　第１問 **問題 1**

論文直前答練第２回　第１問 **問題 1**　　論文直前答練第３回　第１問 **問題 1**

問題 2

論文基礎答練第１回　第１問 **問題 1**　　論文応用答練第１回　第１問 **問題 1**

論文直前答練第１回　第１問 **問題 1**　　論文直前答練第２回　第１問 **問題 1**

問題 3

論文基礎答練第１回　第２問 **問題 3**

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

問題 1

問 1　2017年度を基準時とした2020年度のラスパイレス価格指数は，以下のように求められる。

$$\frac{6480\times46.61+112200\times180.00+6414\times313.96+25400\times964.72}{9930\times46.61+107900\times180.00+6175\times313.96+24430\times964.72}\times100=\frac{47015660.24}{45391649.9}\times100=103.5777\cdots\fallingdotseq103.58$$

同様に，2017年度を基準時とした2021年度のラスパイレス価格指数は，以下のように求められる。

$$\frac{15360\times46.61+128100\times180.00+6175\times313.96+25180\times964.72}{9930\times46.61+107900\times180.00+6175\times313.96+24430\times964.72}\times100=\frac{50004282.2}{45391649.9}\times100=110.1618\cdots\fallingdotseq110.16$$

問 2　2017年度を基準時とした2020年度のパーシェ価格指数は，以下のように求められる。

$$\frac{6480\times38.15+112200\times158.69+6414\times282.62+25400\times913.61}{9930\times38.15+107900\times158.69+6175\times282.62+24430\times913.61}\times100=\frac{43070648.68}{41566151.3}\times100=103.6195\cdots\fallingdotseq103.62$$

同様に，2017年度を基準時とした2021年度のパーシェ価格指数は，以下のように求められる。

$$\frac{15360\times42.75+128100\times158.08+6175\times295.44+25180\times931.94}{9930\times42.75+107900\times158.08+6175\times295.44+24430\times931.94}\times100=\frac{46197279.2}{42072975.7}\times100=109.8027\cdots\fallingdotseq109.80$$

問 3　解答を参照のこと。

問題 2

問 1　解答を参照のこと。

問 2　解答を参照のこと。

問 3　貯蓄0万円の世帯の(全体に対する)割合を x %とすると，貯蓄保有世帯の(全体に対する)割合は(100－ x)%であるから，貯蓄0万円の世帯を除いた貯蓄保有世帯の中央値は，貯蓄0万円の世帯を含めた全体の上側(50－0.5 x)%点（ないしは下側(50＋0.5 x)%点）に対応する。貯蓄0万円の世帯を含めた全体の上側50%点（下側50%点）との間に，1.98%の世帯が入っていると推計できることから，

$$50-(50-0.5x)=1.98$$

より，

$$x=3.96\quad\rightarrow\quad3.96\%$$

と求められる。

問 4　貯蓄0万円の世帯の割合は3.96%，貯蓄保有世帯の割合は96.04%であるから，貯蓄0万円の世帯を含めた(全体の)平均値は，以下のように求められる。

$$\text{平均値}=0.0396\times0+0.9604\times1454=1396.4216\fallingdotseq1396\text{（万円）}$$

問題 3

問 1　景品の入っている箱 X が，参加者が最初に選択した箱 a である確率は $\frac{1}{4}$ である。その条件の下で，司会者は b_1，b_2，b_3 の中から開ける箱を均等な確率で選ぶことになるから，司会者が開ける箱 Y が b_2 である条件付確率は $\frac{1}{3}$ である。したがって，求める同時確率 $\Pr(X=a,\ Y=b_2)$ は，

$$\Pr(X=a,\ Y=b_2)=\frac{1}{4}\times\frac{1}{3}=\frac{1}{12}$$

と計算できる。

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

問 2　景品の入っている箱 X が b_1 である確率は $\frac{1}{4}$ である。その条件の下では，司会者は b_2，b_3 の中から開ける箱を均等な確率で選ぶことになるから，司会者が開ける箱 Y が b_2 である条件付確率は $\frac{1}{2}$ である。したがって，求める同時確率 $\Pr(X=b_1,\ Y=b_2)$ は，

$$\Pr(X=b_1,\ Y=b_2)=\frac{1}{4}\times\frac{1}{2}=\frac{1}{8}$$

と計算できる。

問 3　前問までと同様にして，同時確率 $\Pr(X=b_2,\ Y=b_2)=0$，$\Pr(X=b_3,\ Y=b_2)=\frac{1}{8}$ となるから，司会者が開ける箱 Y が b_2 である周辺確率 $\Pr(Y=b_2)$ は，排反事象の同時確率を足し合わせることにより，

$$\Pr(Y=b_2)=\Pr(X=a,\ Y=b_2)+\Pr(X=b_1,\ Y=b_2)+\Pr(X=b_2,\ Y=b_2)+\Pr(X=b_3,\ Y=b_2)$$

$$=\frac{1}{12}+\frac{1}{8}+0+\frac{1}{8}=\frac{1}{3}$$

と計算できる。

問 4　解答を参照のこと。

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

第8問　答案用紙<1>
(統　計　学)

問題 1

問 1

下限	上限
$\hat{p}_A - 1.96\sqrt{\dfrac{\hat{p}_A(1-\hat{p}_A)}{n_A}}$	$\hat{p}_A + 1.96\sqrt{\dfrac{\hat{p}_A(1-\hat{p}_A)}{n_A}}$

問 2

信頼区間の幅の上限	標本の大きさ
$\dfrac{1.96}{\sqrt{n_A}}$	385　人以上

問 3

278　人以上

問 4

下限	上限
−0.01	0.11

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

第8問　答案用紙＜2＞
(統　計　学)

問題 2

問 1

検定統計量 T	棄却域
$T=\dfrac{\bar{D}}{\dfrac{S_D}{\sqrt{15}}}$	Tの値が1.761を上回っていれば，帰無仮説を棄却する。

問 2

Tの値をtとすると，$t=\dfrac{0.07}{\dfrac{0.18}{\sqrt{15}}}\fallingdotseq 1.51$と求められる。このため，帰無仮説は採択される。

問 3

試行回数15回，成功確率$\dfrac{1}{2}$の二項分布

問 4

W の値	P 値
11	0.035

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

第8問 答案用紙<3>
(統 計 学)

問題 3

問 1

(1)	(2)
$\dfrac{\sum_{i=1}^{24} x_i y_i}{\sum_{i=1}^{24} x_i^2}$	76.58

(3)	
予測値 1148.71 km/s	予測誤差 265.64 km/s

問 2

(ア)	(イ)	(ウ)	(エ)	(オ)	(カ)
(e)	(f)	(d)	(c)	(b)	(g)

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

【解答への道】

Ⅰ　合格ライン

問題 1

比率の区間推定および比率の差の区間推定に関する問題である。素点でみて６割程度の正答率が望まれる。

問題 2

平均の差の検定，符号検定，および，P値に関する問題である。素点でみて６割程度の正答率が望まれる。

問題 3

定数項のない単回帰モデル，および，連続分布の確率密度関数の形状と離散分布の確率関数の形状に関する問題である。素点でみて５割程度の正答率が望まれる。

全体として，第８問の合格ラインは，素点でみて６割弱程度と考えられる。

Ⅱ　答練との対応関係

問題 1

論文基礎答練第３回　第２問 **問題 2**　　論文応用答練第２回　第２問 **問題 1**

問題 2

論文応用答練第２回　第１問 **問題 3**　　論文応用答練第２回　第２問 **問題 2**

論文直前答練第４回　第２問 **問題 1**　　全国公開模試第２回　第８問 **問題 3**

問題 3

論文応用答練第１回　第１問 **問題 3**　　論文応用答練第２回　第１問 **問題 1**

論文直前答練第２回　第２問 **問題 1**　　論文直前答練第４回　第１問 **問題 3**

論文直前答練第４回　第２問 **問題 2**　　全国公開模試第２回　第７問 **問題 3**

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

問題 1

問 1　地域Aの母集団は100万人であるが，母集団が標本の大きさn_Aに対して十分に大きい場合，この母集団は無限母集団とみなすことができる。このとき，標本の大きさn_Aが大きければ，標本に基づく地域Aのブランド認知率$\hat{p}_A$は，中心極限定理により，平均p_A，分散$\dfrac{p_A(1-p_A)}{n_A}$の正規分布に従う。このため，

$$\Pr\left(-1.96 \leqq \frac{\hat{p}_A - p_A}{\sqrt{\dfrac{p_A(1-p_A)}{n_A}}} \leqq 1.96\right) = 0.95$$

となる。ここで，分散$\dfrac{p_A(1-p_A)}{n_A}$に含まれるp_Aを$\hat{p}_A$に置き換えた，

$$-1.96 \leqq \frac{\hat{p}_A - p_A}{\sqrt{\dfrac{\hat{p}_A(1-\hat{p}_A)}{n_A}}} \leqq 1.96$$

をp_Aについて解くことにより，p_Aの信頼係数95％の信頼区間は，近似的に，

$$\hat{p}_A - 1.96\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A}} \leqq p_A \leqq \hat{p}_A + 1.96\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A}}$$

と求められる。

問 2　p_Aの信頼係数95％の信頼区間の幅は，

$$\hat{p}_A + 1.96\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A}} - \left(\hat{p}_A - 1.96\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A}}\right) = 2\times 1.96\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A}}$$

となる。この幅が最大になるのは$\hat{p}_A = \dfrac{1}{2}$のときである。このため，p_Aの信頼係数95％の信頼区間の幅の上限は，

$$2\times 1.96\sqrt{\frac{\frac{1}{2}\times\left(1-\frac{1}{2}\right)}{n_A}} = \frac{1.96}{\sqrt{n_A}}$$

と示される。さらに，この上限が0.1以下となるためには，

$$\frac{1.96}{\sqrt{n_A}} \leqq 0.1 \Leftrightarrow \sqrt{n_A} \geqq \frac{1.96}{0.1}$$

より，

$$n_A \geqq \left(\frac{1.96}{0.1}\right)^2 = 384.16$$

と求められるので，標本の大きさを385人以上とすればよい。

問 3　地域Bの母集団は1,000人であるので，有限母集団と捉えられる。このため，p_Bの信頼係数95％の信頼区間の幅の上限が0.1以下となるためには，有限母集団修正を用いた，

$$2\times 1.96\sqrt{\frac{1000-n_B}{1000-1}\times\frac{\frac{1}{2}\times\left(1-\frac{1}{2}\right)}{n_B}} \leqq 0.1$$

をn_Bについて解くと，

$$n_B \geqq 277.7408\cdots$$

と求められるので，標本の大きさを278人以上とすればよい。

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

問 4　地域AとCの母集団は，ともに100万人であるが，母集団が標本の大きさ400人に対して十分に大きいため，これらの母集団は無限母集団とみなすことができる。標本の大きさが大きいので，$\hat{p}_A-\hat{p}_C$ は，中心極限定理により，平均 p_A-p_C，分散 $\frac{p_A(1-p_A)}{400}+\frac{p_C(1-p_C)}{400}$ の正規分布に従う。このため，

$$\Pr\left(-1.645\leqq\frac{\hat{p}_A-\hat{p}_C-(p_A-p_C)}{\sqrt{\frac{p_A(1-p_A)}{400}+\frac{p_C(1-p_C)}{400}}}\leqq1.645\right)=0.9$$

となる。ここで，分散 $\frac{p_A(1-p_A)}{400}+\frac{p_C(1-p_C)}{400}$ に含まれるp_Aとp_Cを $\hat{p}_A$ と $\hat{p}_C$ に置き換えた，

$$-1.645\leqq\frac{\hat{p}_A-\hat{p}_C-(p_A-p_C)}{\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{400}+\frac{\hat{p}_C(1-\hat{p}_C)}{400}}}\leqq1.645$$

を p_A-p_C について解くことにより，p_A-p_C の信頼係数90%の信頼区間は，近似的に，

$$\hat{p}_A-\hat{p}_C-1.645\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{400}+\frac{\hat{p}_C(1-\hat{p}_C)}{400}}\leqq p_A-p_C\leqq\hat{p}_A-\hat{p}_C+1.645\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{400}+\frac{\hat{p}_C(1-\hat{p}_C)}{400}}$$

と示される。$\hat{p}_A=\frac{180}{400}$，$\hat{p}_C=\frac{160}{400}$ であるので，p_A-p_C の信頼係数90%の信頼区間の下限と上限は，

$$\text{下限}=0.45-0.4-1.645\sqrt{\frac{0.45(1-0.45)}{400}+\frac{0.4(1-0.4)}{400}}=-0.007427\cdots\fallingdotseq-0.01$$

$$\text{上限}=0.45-0.4+1.645\sqrt{\frac{0.45(1-0.45)}{400}+\frac{0.4(1-0.4)}{400}}=0.107427\cdots\fallingdotseq0.11$$

と求められる。

問題 2

問 1　検定統計量

$$T=\frac{\bar{D}}{\frac{S_D}{\sqrt{15}}}$$

は，帰無仮説のもとで自由度14のt分布に従う。自由度14のt分布の上側5%点の値は1.761であるので，有意水準を5%とした場合，Tの値が1.761を上回っていれば，帰無仮説を棄却する。

問 2　解答を参照のこと。

問 3　解答を参照のこと。

問 4　統計量Wは，問 3 より，試行回数15回，成功確率$\frac{1}{2}$の二項分布に従う。このため，Wの平均は$15\times\frac{1}{2}=7.5$，分散は$15\times\frac{1}{2}\times\frac{1}{2}=3.75$となる。$W=11$に対するP値は，$W$が大きいときに$H_0$を棄却する検定において，正規近似を用いると，

$$\Pr(W\geqq11)=\Pr\left(\frac{W-7.5}{\sqrt{3.75}}\geqq\frac{11-7.5}{\sqrt{3.75}}\right)=\Pr(Z\geqq1.807\cdots)\fallingdotseq0.035$$

と求められる。ただし，Zは標準正規分布に従う確率変数をあらわす。なお，この問では端数処理の指定がないため，標準正規分布表より$Z=1.80$のときの上側確率が0.0359，$Z=1.81$のときの上側確率が0.0351であることより，解答を0.035としている。

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

問題 3

問 1　(1) 回帰係数βの最小二乗推定値$\hat{\beta}$を

$$Q=\sum_{i=1}^{24}(y_i-\beta x_i)^2$$

を最小にするβとするとき，Qをβで微分してゼロとおくと，

$$\frac{dQ}{d\beta}=0 \quad \Leftrightarrow \quad 2\times\sum_{i=1}^{24}\{(y_i-\beta x_i)\times(-x_i)\}=0$$

となるので，

$$\sum_{i=1}^{24}\{(y_i-\beta x_i)\times x_i\}=0 \quad \Leftrightarrow \quad \sum_{i=1}^{24}x_iy_i-\beta\times\sum_{i=1}^{24}x_i^2=0$$

より，回帰係数βの推定値$\hat{\beta}$は，

$$\hat{\beta}=\frac{\sum_{i=1}^{24}x_i\,y_i}{\sum_{i=1}^{24}x_i^2}$$

と示される。

(2) 最小二乗推定値$\hat{\beta}$の値は，

$$\hat{\beta}=\frac{\sum_{i=1}^{24}x_i\,y_i}{\sum_{i=1}^{24}x_i^2}=\frac{326629.08}{4265.15}=76.5809\cdots\fallingdotseq 76.58$$

と求められる。

(3) 距離が15Mpcである星雲の後退速度の予測値は，

予測値＝76.5809…×15＝1148.7136…≒1148.71（km/s）

と求められる。さらに，予測誤差は，

$$\hat{\beta}x_0+\varepsilon$$

の標準偏差であるので，

$$\text{予測誤差}=\sqrt{\mathrm{V}\left(\hat{\beta}x_0+\varepsilon\right)}=\sqrt{x_0^2\,\mathrm{V}\left(\hat{\beta}\right)+\mathrm{V}(\varepsilon)}$$

より求められる。ここで，V($\hat{\beta}$)の値は，出力結果にあるStd.Errorの値3.965の2乗となる。また，誤差の分散σ^2の値は，出力結果にあるResidual standard errorの値258.9の2乗となる。これらより，予測誤差は，

$$\text{予測誤差}=\sqrt{15^2\times3.965^2+258.9^2}=265.6435\cdots\fallingdotseq 265.64\ (\text{km/s})$$

と求められる。

問 2　（ア）～（ウ）は連続分布であるが，問題に与えられている分布のうち，連続分布は，

(a)　平均5，分散5の正規分布

(d)　パラメータα＝2，β＝3のベータ分布

(e)　形状（shape）パラメータ5，尺度（scale）パラメータ1のガンマ分布

(f)　平均5の指数分布

(h)　区間[0，1]の一様分布

の5つの分布である。このうち，(a)と(h)は，ヒストグラムの形状および定義域から，（ア）～（ウ）に含まれないと判断できる。指数分布の特徴は，小さい値ほど出やすいことにあるので，（イ）は(f)と判断できる。また，定義域から（ウ）は(d)と判断できる。これらより，（ア）は（e）となる。

この解答速報の著作権はＴＡＣ(株)のものであり，無断転載・転用を禁じます。

一方，（エ）～（カ）は離散分布であるが，問題に与えられている分布のうち，離散分布は，

(b)　平均5のポアソン分布

(c)　試行回数12，生起確率$\frac{5}{12}$の2項分布

(g)　状態18，成功状態数12，非復元抽出数12の超幾何分布

の3つの分布である。このうち，（エ）は定義域から(c)と判断できる。また，（カ）も定義域から(g)と判断できる。これらより，（オ）は（b）となる。