

第7問 答案用紙<1> (統計学)

問題 1

ア	イ	ウ
500,000	360,000	30

エ
360,000

オ
416,000

カ	キ
0.25	0.62

(それぞれ四捨五入して小数点以下 2 ケタ)

ク	ケ
0.31	0.57

(それぞれ四捨五入して小数点以下 2 ケタ)

第7問 答案用紙<2>

(統計学)

問題2

(1)

ア	イ	ウ	エ	オ
0.10	0.30	0.40	0.15	0.05

(それぞれ小数点以下2ケタ)

(2)

カ	キ	ク	ケ	コ
0.025	0.210	0.420	0.220	0.125

(それぞれ小数点以下3ケタ)

(3) ジニ係数 = 0.24

(四捨五入して小数点以下2ケタ)

(計算の詳細)

累積相対度数と累積所得比率は、以下の表のように整理される。

階級値	150	420	630	880	1500
累積相対度数	0.1	0.4	0.8	0.95	1
累積所得比率	0.025	0.235	0.655	0.875	1

この表より、ジニ係数は、つぎのように求められる。

$$\begin{aligned}
 \text{ジニ係数} &= 1 - 2 \times \left\{ \frac{1}{2} \times 0.1 \times 0.025 + \frac{1}{2} \times (0.4 - 0.1) \times (0.025 + 0.235) \right. \\
 &\quad + \frac{1}{2} \times (0.8 - 0.4) \times (0.235 + 0.655) + \frac{1}{2} \times (0.95 - 0.8) \times (0.655 + 0.875) \\
 &\quad \left. + \frac{1}{2} \times (1 - 0.95) \times (0.875 + 1) \right\} \\
 &= 0.24025 \div 0.24
 \end{aligned}$$

(4)

サ	シ	ス	セ
ヒストグラム	面積	右	算術平均

第7問 答案用紙<3> (統計学)

問題3

(1) $V(X) =$

$$\frac{3}{4}$$

(2) $\text{Cov}(Y, Z) =$

$$-6$$

(3) $\text{Cor}(Z, W) =$

$$0$$

(4) $\Pr(Z = 8|W = 2) =$

$$\frac{3}{4}$$

【解答への道】

I 合格ライン

問題 1

確率の分野から、期待値や条件付確率の計算を中心とした出題である。基本的な計算問題であり、合格ラインは6割程度と思われるが、TACの答練での出題があるため、TAC生ならば素点で見て6～7割程度の正答率が望まれるであろう。

問題 2

記述統計の分野から、度数分布表、ジニ係数、ヒストグラムに関する出題である。ジニ係数の計算がやや煩雑であるが、全体的には基本的な計算問題であり、素点で見て6～7割程度の正答率が望まれる。

問題 3

2つの独立な確率変数と変数変換に関する標準的な問題である。合格ラインは素点で見て5割程度と思われる。

全体として、第7問の合格ラインは、素点で見て60%を多少超える程度と考えられる。

II 答練との対応関係

問題 1

基礎答練第1回 第2問	問題 3	基礎答練第2回 第1問	問題 1
基礎答練第4回 第1問	問題 1		

問題 2

基礎答練第1回 第1問	問題 1	問題 2	
応用答練第1回 第1問	問題 2	直前答練第2回 第1問	問題 1
直前答練第3回 第1問	問題 1	直前答練第4回 第1問	問題 1

問題 3

基礎答練第1回 第1問	問題 3	応用答練第1回 第1問	問題 3
直前答練第2回 第1問	問題 3	直前答練第4回 第1問	問題 2

問題 1

1. について

《空欄 ア ～ ウ について》

天気が晴れのとき、曇りのときの売上金額の平均はそれぞれ、(価格)×(数量)の合計を計算することにより、以下のように求められる。

晴れのときの売上金額の平均 $=300 \times 100 + 400 \times 200 + 500 \times 300 + 600 \times 400 = 500000$ (円) (空欄 ア の答)

曇りのときの売上金額の平均 $=300 \times 200 + 400 \times 200 + 500 \times 200 + 600 \times 200 = 360000$ (円) (空欄 イ の答)

空欄 ウ の数量を x とすると、天気が雨の時の売上金額の平均についても上記と同様の関係式が成立することから、
雨のときの売上金額の平均 $=300 \times 60 + 400 \times 60 + 500 \times 40 + 600 \times x = 80000$ (円)

より、

$$x = 30 \text{ (個)} \quad \text{(空欄 ウ の答)}$$

と求められる。

(1) 天気が晴れ、または、曇りであるとわかっているとき、売上金額は360000(円) (空欄 エ の答) 以上であると期待される。本問で問われているのは、いわゆる「期待値」ではなく、期待される売上金額の平均の最低ライン (すなわち、500000円と360000円のうち、小さい方の金額) であることに注意しよう。

(2) 売上金額の期待値は、(確率)×(売上金額の平均)の合計をとることにより、以下のように求められる。

$$\text{売上金額の期待値} = 0.4 \times 500000 + 0.6 \times 360000 + 0.0 \times 80000 = 416000 \text{ (円)} \quad \text{(空欄 オ の答)}$$

2. について

(1)

1) 弁当が1個売れる状況について、同時確率と周辺確率を考える。天気が晴れで、かつ400円の弁当が1つ売れる同時確率は、

$$\frac{1}{3} \times 0.2 = \frac{1}{15}$$

である。また、400円の弁当が1つ売れる周辺確率は、

$$\frac{1}{3} \times 0.2 + \frac{1}{3} \times 0.3 + \frac{1}{3} \times 0.3 = \frac{4}{15}$$

である。このため、ベイズの定理を用いると、求める条件付確率は、

$$\frac{\left(\frac{1}{15}\right)}{\left(\frac{4}{15}\right)} = \frac{1}{4} = 0.25 \quad \text{(空欄 カ の答)}$$

と計算できる。

2) 弁当が3個売れる状況について、まずは条件付確率を考える。500円の弁当が2つ、600円の弁当が1つ売れる確率は、天気が晴れ、曇り、雨のそれぞれの条件下で、以下のように、2項分布の確率として求められる。

天気が晴れの下、500円の弁当が2つ、600円の弁当が1つ売れる条件付確率 $= {}_3C_2 \cdot 0.3^2 \cdot 0.4 = 0.108$

天気が曇りの下、500円の弁当が2つ、600円の弁当が1つ売れる条件付確率 $= {}_3C_2 \cdot 0.3^2 \cdot 0.2 = 0.054$

天気が雨の下、500円の弁当が2つ、600円の弁当が1つ売れる条件付確率 $= {}_3C_2 \cdot 0.2^2 \cdot 0.1 = 0.012$

つぎに、同時確率と周辺確率を考える。天気が晴れで、かつ、500円の弁当が2つと600円の弁当が1つ売れる同時確率は、

$$\frac{1}{3} \times 0.108 = 0.036$$

である。また、500円の弁当が2つと600円の弁当が1つ売れる周辺確率は、

$$\frac{1}{3} \times 0.108 + \frac{1}{3} \times 0.054 + \frac{1}{3} \times 0.012 = 0.058$$

である。このため、ベイズの定理を用いると、求める条件付確率は、以下のように計算できる。

$$\frac{0.036}{0.058} = 0.6206 \cdots \div 0.62 \quad \text{(空欄 キ の答)}$$

この解答速報の著作権はT A C (株)のものであり，無断転載・転用を禁じます。

(2)

- 1) (1)の場合と同じように，弁当が1個売れる状況について，同時確率と周辺確率を考える。天気が晴れで，かつ400円の弁当が1つ売れる同時確率は，

$$0.4 \times 0.2 = 0.08$$

である。また，400円の弁当が1つ売れる周辺確率は，

$$0.4 \times 0.2 + 0.6 \times 0.3 + 0.0 \times 0.3 = 0.26$$

である。このため，ベイズの定理を用いると，求める条件付確率は，以下のように計算できる。

$$\frac{0.08}{0.26} = 0.3076 \cdots \approx 0.31 \text{ (空欄 } \boxed{\text{ク}} \text{ の答)}$$

- 2) 弁当が3個売れる状況についての条件付確率は，(1)で求めたとおりである。

同時確率と周辺確率を考える。天気が晴れで，かつ，500円の弁当が2つと600円の弁当が1つ売れる同時確率は，

$$0.4 \times 0.108 = 0.0432$$

である。また，500円の弁当が2つと600円の弁当が1つ売れる周辺確率は，

$$0.4 \times 0.108 + 0.6 \times 0.054 + 0.0 \times 0.012 = 0.0756$$

である。このため，ベイズの定理を用いると，求める条件付確率は，

$$\frac{0.0432}{0.0756} = 0.5714 \cdots \approx 0.57 \text{ (空欄 } \boxed{\text{ケ}} \text{ の答)}$$

と計算できる。

問題 2

- (1) ある階級の累積度数が，その階級の度数とその前の階級の累積度数の合計となることに留意すると，

$$250\text{万円未満の階級の度数} = 400$$

$$750\text{万円以上} \sim 1000\text{万円未満の階級の度数} = 600$$

と，空欄を埋められる。また，250万円以上～500万円未満の階級の累積度数＝1600が，総度数の40%に対応することに注目すると，総度数が4000であることがわかるから，

$$1000\text{万円以上の階級の度数} = 200$$

と，空欄を埋められる。

このため，階級ごとの相対度数は，その階級の度数を総度数で割ることにより，以下のように求められる。

$$250\text{万円未満の階級の相対度数} = \frac{400}{4000} = 0.10 \text{ (空欄 } \boxed{\text{ア}} \text{ の答)}$$

$$250\text{万円以上} \sim 500\text{万円未満の階級の相対度数} = \frac{1200}{4000} = 0.30 \text{ (空欄 } \boxed{\text{イ}} \text{ の答)}$$

$$500\text{万円以上} \sim 750\text{万円未満の階級の相対度数} = \frac{1600}{4000} = 0.40 \text{ (空欄 } \boxed{\text{ウ}} \text{ の答)}$$

$$750\text{万円以上} \sim 1000\text{万円未満の階級の相対度数} = \frac{600}{4000} = 0.15 \text{ (空欄 } \boxed{\text{エ}} \text{ の答)}$$

$$1000\text{万円以上の階級の相対度数} = \frac{200}{4000} = 0.05 \text{ (空欄 } \boxed{\text{オ}} \text{ の答)}$$

- (2) 各階級の所得合計と全体の所得合計は，(階級値)×(度数)を考えることにより，以下のように求められる。

$$250\text{万円未満の階級の所得合計} = 150 \times 400 = 60000$$

$$250\text{万円以上} \sim 500\text{万円未満の階級の所得合計} = 420 \times 1200 = 504000$$

$$500\text{万円以上} \sim 750\text{万円未満の階級の所得合計} = 630 \times 1600 = 1008000$$

$$750\text{万円以上} \sim 1000\text{万円未満の階級の所得合計} = 880 \times 600 = 528000$$

$$1000\text{万円以上の階級の所得合計} = 1500 \times 200 = 300000$$

$$\text{全体の所得合計} = 60000 + 504000 + 1008000 + 528000 + 300000 = 2400000$$

このため，各階級の所得比率は，以下のように求められる。

この解答速報の著作権はT A C (株)のものであり、無断転載・転用を禁じます。

$$250万円未満の階級の所得比率 = \frac{60000}{2400000} = 0.025 \text{ (空欄「カ」の答)}$$

$$250万円以上～500万円未満の階級の所得比率 = \frac{504000}{2400000} = 0.210 \text{ (空欄「キ」の答)}$$

$$500万円以上～750万円未満の階級の所得比率 = \frac{1008000}{2400000} = 0.420 \text{ (空欄「ク」の答)}$$

$$750万円以上～1000万円未満の階級の所得比率 = \frac{528000}{2400000} = 0.220 \text{ (空欄「ケ」の答)}$$

$$1000万円以上の階級の所得比率 = \frac{300000}{2400000} = 0.125 \text{ (空欄「コ」の答)}$$

(3) 解答を参照のこと。

(4) 一般に、所得や貯蓄の分布は、右に裾を引いた形（右に歪んだ分布）になることが知られている。右に歪んだ分布においては、分布の中心を示す代表値に関して、

$$\text{モード(最頻値)} < \text{メディアン(中央値)} < \text{算術平均}$$

という関係が成立する。

問題 3

(1) 確率変数 X の期待値 $E(X)$ は、

$$E(X) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{12}{8} = \frac{3}{2}$$

と求められる。また、 X^2 の期待値 $E(X^2)$ は、

$$E(X^2) = 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8} = \frac{24}{8} = 3$$

と求められる。これらにより、 X の分散 $V(X)$ は、以下のように求められる。

$$V(X) = E(X^2) - \{E(X)\}^2 = 3 - \left(\frac{3}{2}\right)^2 = \frac{3}{4} \text{ (答)}$$

(2) 確率変数 Y と Z の共分散 $\text{Cov}(Y, Z)$ は、

$$\text{Cov}(Y, Z) = E(YZ) - E(Y) \times E(Z)$$

とあらわされる。ここで、確率変数 Y の期待値 $E(Y)$ は、

$$E(Y) = (-2) \times \frac{1}{5} + (-1) \times \frac{1}{5} + 0 \times \frac{1}{5} + 1 \times \frac{1}{5} + 2 \times \frac{1}{5} = 0$$

と求められる。また、確率変数 Z の期待値 $E(Z)$ は、

$$E(Z) = E(2X - 3Y) = 2E(X) - 3E(Y) = 2 \times \frac{3}{2} - 3 \times 0 = 3$$

と求められ、 YZ の期待値 $E(YZ)$ は、

$$E(YZ) = E\{Y(2X - 3Y)\} = 2E(XY) - 3E(Y^2)$$

とあらわされる。確率変数 X と Y は互いに独立であるから、

$$E(XY) = E(X) \times E(Y) = \frac{3}{2} \times 0 = 0$$

と求められ、 Y^2 の期待値 $E(Y^2)$ は、

$$E(Y^2) = (-2)^2 \times \frac{1}{5} + (-1)^2 \times \frac{1}{5} + 0^2 \times \frac{1}{5} + 1^2 \times \frac{1}{5} + 2^2 \times \frac{1}{5} = 2$$

と求められる。これらにより、 YZ の期待値 $E(YZ)$ は、

$$E(YZ) = 2E(XY) - 3E(Y^2) = 2 \times 0 - 3 \times 2 = -6$$

と求められる。以上のことより、 Y と Z の共分散 $\text{Cov}(Y, Z)$ は、

$$\text{Cov}(Y, Z) = E(YZ) - E(Y) \times E(Z) = -6 - 0 \times 3 = -6 \text{ (答)}$$

と求められる。

この解答速報の著作権はT A C (株)のものであり、無断転載・転用を禁じます。

(3) 確率変数 Z と W の共分散 $\text{Cov}(Z, W)$ は、

$$\text{Cov}(Z, W) = E(ZW) - E(Z) \times E(W)$$

とあらわされる。ここで、確率変数 W の期待値 $E(W)$ は、

$$E(W) = E(4X + Y) = 4E(X) + E(Y) = 4 \times \frac{3}{2} + 0 = 6$$

と求められる。また、 ZW の期待値 $E(ZW)$ は、

$$\begin{aligned} E(ZW) &= E\{(2X - 3Y)(4X + Y)\} = E(8X^2 - 10XY - 3Y^2) \\ &= 8E(X^2) - 10E(XY) - 3E(Y^2) = 8 \times 3 - 10 \times 0 - 3 \times 2 = 18 \end{aligned}$$

と求められる。このため、 $\text{Cov}(Z, W)$ は、

$$\text{Cov}(Z, W) = E(ZW) - E(Z) \times E(W) = 18 - 3 \times 6 = 0$$

と求められる。 Z と W の共分散が0であるから、相関係数についても、 $\text{Cor}(Z, W) = 0$ (答) となる。

(4) X と Y のとり得る値に注目すると、 $W=2$ となるのは、 $(x, y) = (0, 2)$ または $(x, y) = (1, -2)$ の場合であることがわかる。 $W=2$ となる確率 (周辺確率) は、

$$\frac{1}{8} \times \frac{1}{5} + \frac{3}{8} \times \frac{1}{5} = \frac{4}{40} = \frac{1}{10}$$

である。このうち、 $Z=8$ となるのは、 $(x, y) = (1, -2)$ の場合のみであり、その確率 (同時確率) は、

$$\frac{3}{8} \times \frac{1}{5} = \frac{3}{40}$$

である。このため、条件付確率 $\text{Pr}(Z=8 | W=2)$ は、以下のように求められる。

$$\text{Pr}(Z=8 | W=2) = \frac{\left(\frac{3}{40}\right)}{\left(\frac{1}{10}\right)} = \frac{3}{4} \quad (\text{答})$$

第8問 答案用紙<1> (統計学)

問題 1

(1) $\hat{\alpha} =$

$$\frac{1}{5} \sum_{h=1}^5 y_h - \hat{\beta} \frac{1}{5} \sum_{h=1}^5 x_h$$

$\hat{\beta} =$

$$\frac{\sum_{h=1}^5 x_h y_h - \frac{1}{5} \sum_{h=1}^5 x_h \sum_{h=1}^5 y_h}{\sum_{h=1}^5 x_h^2 - \frac{1}{5} \left(\sum_{h=1}^5 x_h \right)^2}$$

(2) $\hat{\alpha} =$

-20.00

$\hat{\beta} =$

0.12

(それぞれ四捨五入して小数点以下2ケタ)

(3) $\hat{\sigma}^2 =$

53.333

(四捨五入して小数点以下3ケタ)

(4) 決定係数 =

0.90

(四捨五入して小数点以下2ケタ)

(5) 予測値 =

46

(四捨五入して万円単位まで)

第8問 答案用紙<2> (統計学)

問題2

(1) (仮説検定の詳細と検定結果)

帰無仮説を $\sigma_A^2=25$ ，対立仮説を $\sigma_A^2<25$ と設定する。このとき，検定統計量として，

$$V = \frac{30S_A^2}{\sigma_A^2}$$

をもちいると，この検定統計量 V は，帰無仮説のもとで，自由度30のカイ2乗分布に従う。自由度30のカイ2乗分布の上側95%点の値は18.49であるため，有意水準を5%とすると，この片側検定の棄却域は $V<18.49$ となる。一方，検定統計量 V の値は，帰無仮説のもとで， $V=21.168$ と求められる。これらより，帰無仮説は採択されるため，検定対象である主張は統計的に有意ではない。

(2) (仮説検定の詳細と検定結果)

帰無仮説を $\sigma_A^2=\sigma_B^2$ ，対立仮説を $\sigma_A^2<\sigma_B^2$ と設定する。このとき，検定統計量として，

$$W = \frac{S_B^2}{S_A^2}$$

をもちいると，この検定統計量 W は，帰無仮説のもとで，自由度(15, 30)の F 分布に従う。自由度(15, 30)の F 分布の上側5%点の値は2.015であるため，有意水準を5%とすると，この片側検定の棄却域は $W<2.015$ となる。一方，検定統計量 W の値は $W=1.9070\cdots$ と求められる。これらより，帰無仮説は採択されるため，検定対象である主張は統計的に有意ではない。

(3) 信頼区間の下限＝

11.264

信頼区間の上限＝

31.519

(それぞれ四捨五入して小数点以下3ケタ)

第8問 答案用紙<3> (統計学)

問題3

ア	イ	ウ	エ
μ_i	σ^2	正規	μ_i

オ	カ	キ
$\frac{\sigma^2}{n_i}$	μ	$\frac{\sigma^2}{n}$

ク	ケ	コ	サ
$n-l$	カイ2乗	0	$l-1$

シ	ス	セ	ソ
S_B	$l-1$	S_W	$n-l$

タ	チ	ツ
$l-1$	$n-l$	F

【解答への道】

I 合格ライン

問題 1

単回帰分析についての出題である。基本的な内容の問題であり，素点で見て6～7割程度の正答率が望まれる。

問題 2

分散に関する区間推定，および，分散・分散比の検定についての出題である。いずれも基本的な内容であり，素点で見て6～7割程度の正答率が望まれる。

問題 3

分散分析についての出題である。標本平均の標本分布，および，平均からの偏差の平方和とカイ2乗分布の関係に関する基本的な理論を十分に理解していれば，部分点を稼ぎ出すことができるので，素点で見て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，素点で見て60%を多少超える程度と考えられる。

II 答練との対応関係

問題 1

全答練第2回 第8問 問題 2 直前答練第1回 第2問 問題 3

問題 2

直前答練第1回 第2問 問題 1 基礎答練第3回 第1問 問題 3

基礎答練第3回 第2問 問題 1

問題 3

基礎答練第4回 第2問 問題 3 全答練第1回 第7問 問題 3 など

全答練第2回 第7問 問題 3 など

問題 1

- (1) 所得 x のデータ x_h ($h=1, \dots, 5$) の分散を $V(x)$, 所得 x のデータ x_h と貯蓄 y のデータ y_h ($h=1, \dots, 5$) の共分散を $Cov(x, y)$ とすると、回帰方程式 $y_h = \alpha + \beta x_h + u_h$ ($h=1, \dots, 5$) の β の最小2乗推定量 $\hat{\beta}$ は、つぎのように示される。

$$\hat{\beta} = \frac{Cov(x, y)}{V(x)} = \frac{\sum_{h=1}^5 \left(x_h - \frac{1}{5} \sum_{h=1}^5 x_h \right) \left(y_h - \frac{1}{5} \sum_{h=1}^5 y_h \right)}{\sum_{h=1}^5 \left(x_h - \frac{1}{5} \sum_{h=1}^5 x_h \right)^2} = \frac{\sum_{h=1}^5 x_h y_h - \frac{1}{5} \left(\sum_{h=1}^5 x_h \right) \left(\sum_{h=1}^5 y_h \right)}{\sum_{h=1}^5 x_h^2 - \frac{1}{5} \left(\sum_{h=1}^5 x_h \right)^2}$$

一方、所得 x のデータ x_h ($h=1, \dots, 5$) の平均を $\bar{x}$, 貯蓄 y のデータ y_h ($h=1, \dots, 5$) の平均を $\bar{y}$ とすると、回帰方程式 $y_h = \alpha + \beta x_h + u_h$ ($h=1, \dots, 5$) の α の最小2乗推定量 $\hat{\alpha}$ は、つぎのように示される。

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x} = \frac{1}{5} \sum_{h=1}^5 y_h - \hat{\beta} \frac{1}{5} \sum_{h=1}^5 x_h$$

- (2) 問題文で与えられたデータより、以下の計算結果を得る。

$$\sum_{h=1}^5 x_h = 2500, \quad \sum_{h=1}^5 x_h^2 = 1350000, \quad \sum_{h=1}^5 y_h = 200, \quad \sum_{h=1}^5 y_h^2 = 9600, \quad \sum_{h=1}^5 x_h y_h = 112000$$

これらの値を(1)で示した式に代入すると、最小2乗推定量 $\hat{\alpha}$ と $\hat{\beta}$ の値を以下のように求めることができる。

$$\begin{aligned} \hat{\beta} &= \frac{\sum_{h=1}^5 x_h y_h - \frac{1}{5} \left(\sum_{h=1}^5 x_h \right) \left(\sum_{h=1}^5 y_h \right)}{\sum_{h=1}^5 x_h^2 - \frac{1}{5} \left(\sum_{h=1}^5 x_h \right)^2} = \frac{112000 - \frac{1}{5} \times 2500 \times 200}{1350000 - \frac{1}{5} \times 2500^2} = 0.12 \\ \hat{\alpha} &= \frac{1}{5} \sum_{h=1}^5 y_h - \hat{\beta} \frac{1}{5} \sum_{h=1}^5 x_h = \frac{200}{5} - \frac{0.12 \times 2500}{5} = -20.00 \end{aligned}$$

- (3) 誤差項 u_h の分散 σ^2 の不偏推定量 $\hat{\sigma}^2$ は以下の式で示すことができる。

$$\hat{\sigma}^2 = \frac{1}{5-2} \sum_{h=1}^5 (y_h - \hat{\alpha} - \hat{\beta} x_h)^2$$

ここで、最小2乗推定量 $\hat{\alpha}$ と $\hat{\beta}$ の値と、問題文で与えられたデータより、

$$\sum_{h=1}^5 (y_h - \hat{\alpha} - \hat{\beta} x_h)^2 = \{(-10)^2 + 6^2 + (-2)^2 + 2^2 + 4^2\} = 160$$

と求められるので、 σ^2 の不偏推定量の値は、つぎのように求められる。

$$\hat{\sigma}^2 = \frac{1}{5-2} \sum_{h=1}^5 (y_h - \hat{\alpha} - \hat{\beta} x_h)^2 = \frac{1}{5-2} \times 160 = 53.3333 \dots \doteq 53.333$$

- (4) 所得 x のデータ x_h と貯蓄 y のデータ y_h ($h=1, \dots, 5$) の相関係数 r_{xy} は、(2)で求めた計算結果をもちいると、つぎのように求められる。

$$\begin{aligned} r_{xy} &= \frac{\frac{1}{5} \sum_{h=1}^5 x_h y_h - \left(\frac{1}{5} \sum_{h=1}^5 x_h \right) \times \left(\frac{1}{5} \sum_{h=1}^5 y_h \right)}{\sqrt{\frac{1}{5} \sum_{h=1}^5 x_h^2 - \left(\frac{1}{5} \sum_{h=1}^5 x_h \right)^2} \times \sqrt{\frac{1}{5} \sum_{h=1}^5 y_h^2 - \left(\frac{1}{5} \sum_{h=1}^5 y_h \right)^2}} \\ &= \frac{\frac{1}{5} \times 112000 - \left(\frac{1}{5} \times 2500 \right) \times \left(\frac{1}{5} \times 200 \right)}{\sqrt{\frac{1}{5} \times 1350000 - \left(\frac{1}{5} \times 2500 \right)^2} \times \sqrt{\frac{1}{5} \times 9600 - \left(\frac{1}{5} \times 200 \right)^2}} \\ &= \frac{2400}{\sqrt{20000} \times \sqrt{320}} \end{aligned}$$

決定係数 R^2 は、この相関係数 r_{xy} を2乗したものに等しいので、つぎのように求められる。

$$R^2 = r_{xy}^2 = \left(\frac{2400}{\sqrt{20000} \times \sqrt{320}} \right)^2 = 0.90$$

- (5) (2)で求めた最小2乗推定量 $\hat{\alpha}$ と $\hat{\beta}$ の値をもとにすると，所得が $x_h=550$ 万円のときの家計の貯蓄の予想値 $\hat{y}_h$ は，つぎのように求められる。

$$\hat{y}_h = \hat{\alpha} + \hat{\beta} x_h = -20.00 + 0.12 \times 550 = 46 \text{ (万円)}$$

問題 2

- (1) 解答を参照のこと。
- (2) 解答を参照のこと。
- (3) メーカーAの電球の耐用時間は，平均 μ ，分散 σ_A^2 の正規分布に従っているので，メーカーAの電球31個を無作為に選んで耐用時間を観察した結果得られる標本(不偏)分散を s_A^2 とすると，

$$\frac{30s_A^2}{\sigma^2}$$

は，自由度30のカイ2乗分布に従う。自由度30のカイ2乗分布の上側2.5%点の値は46.98，自由度30のカイ2乗分布の上側97.5%点の値は16.79であるので，メーカーAの電球の耐用時間の分散 σ_A^2 の95%信頼区間は，標本(不偏)分散が $s_A^2=17.64$ であるとき，

$$16.79 \leq \frac{30 \times 17.64}{\sigma_A^2} \leq 46.98$$

より，

$$\frac{30 \times 17.64}{46.98} \leq \sigma_A^2 \leq \frac{30 \times 17.64}{16.79}$$

と求められる。これより，つぎのような解答を得る。

$$\text{信頼区間の下限} = 11.26436 \cdots \doteq 11.264$$

$$\text{信頼区間の上限} = 31.51876 \cdots \doteq 31.519$$

問題 3

- (1) y_{ij} の平均 $E(y_{ij})$ および分散 $V(y_{ij})$ は，以下のよう求められる。

$$E(y_{ij}) = E(\mu_i + \varepsilon_{ij}) = E(\mu_i) + E(\varepsilon_{ij}) = \mu_i$$

$$V(y_{ij}) = V(\mu_i + \varepsilon_{ij}) = V(\varepsilon_{ij}) = \sigma^2$$

ここで， ε_{ij} は正規分布に従っているので，正規分布の再生性により， y_{ij} も正規分布に従う。

つぎに， $\bar{y}_i$ の平均 $E(\bar{y}_i)$ および分散 $V(\bar{y}_i)$ は，以下のよう求められる。

$$E(\bar{y}_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} E(y_{ij}) = \frac{1}{n_i} \sum_{j=1}^{n_i} \mu_i = \frac{n_i \mu_i}{n_i} = \mu_i$$

$$V(\bar{y}_i) = \frac{1}{n_i^2} \sum_{j=1}^{n_i} V(y_{ij}) = \frac{1}{n_i^2} \sum_{j=1}^{n_i} \sigma^2 = \frac{n_i \sigma^2}{n_i^2} = \frac{\sigma^2}{n_i}$$

ここで， y_{ij} は正規分布に従っているので，正規分布の再生性により， $\bar{y}_i$ も正規分布に従う。

また， $\bar{y}$ の平均 $E(\bar{y})$ は，つぎのように求められる。

$$E(\bar{y}) = E\left(\frac{1}{n} \sum_{i=1}^l n_i \bar{y}_i\right) = \frac{1}{n} \sum_{i=1}^l n_i E(\bar{y}_i) = \frac{1}{n} \sum_{i=1}^l n_i \mu_i$$

ここで，問題文の仮定 $\alpha_i = \mu_i - \mu$ より， $\mu_i = \alpha_i + \mu$ となるので，

$$E(\bar{y}) = \frac{1}{n} \sum_{i=1}^l n_i \mu_i = \frac{1}{n} \sum_{i=1}^l n_i (\alpha_i + \mu) = \frac{1}{n} \sum_{i=1}^l n_i \alpha_i + \frac{1}{n} \sum_{i=1}^l n_i \mu = \frac{1}{n} \sum_{i=1}^l n_i \alpha_i + \frac{\mu}{n} \sum_{i=1}^l n_i$$

となる。さらに，問題文の仮定 $\sum_{i=1}^l n_i \alpha_i = 0$ ， $\sum_{i=1}^l n_i = n$ より，

$$E(\bar{y}) = \frac{1}{n} \sum_{i=1}^l n_i \alpha_i + \frac{\mu}{n} \sum_{i=1}^l n_i = 0 + \frac{\mu}{n} \times n = \mu$$

となる。一方、 $\bar{y}$ の分散 $V(\bar{y})$ は、 $V(\bar{y}_i)=\frac{\sigma^2}{n_i}$ 、および、問題文の仮定 $\sum_{i=1}^l n_i=n$ より、つぎのように求められる。

$$V(\bar{y})=V\left(\frac{1}{n}\sum_{i=1}^l n_i \bar{y}_i\right)=\frac{1}{n^2}\sum_{i=1}^l n_i^2 V(\bar{y}_i)=\frac{1}{n^2}\sum_{i=1}^l \left(n_i^2 \times \frac{\sigma^2}{n_i}\right)=\frac{\sigma^2}{n^2}\sum_{i=1}^l n_i=\frac{\sigma^2}{n^2} \times n=\frac{\sigma^2}{n}$$

ここで、 $\bar{y}_i$ は正規分布に従っているので、正規分布の再生性により、 $\bar{y}$ も正規分布に従う。

(2) y_{ij} は、平均 μ_i 、分散 σ^2 の正規分布に従うので、

$$\frac{1}{\sigma^2}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)^2$$

は、自由度 n_i-1 のカイ2乗分布に従う。このため、カイ2乗分布の再生性より、

$$\frac{S_W}{\sigma^2}=\frac{1}{\sigma^2}\sum_{i=1}^l \sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)^2$$

は、問題文の仮定 $\sum_{i=1}^l n_i=n$ において、自由度 $\sum_{i=1}^l (n_i-1)=\sum_{i=1}^l n_i-l=n-l$ のカイ2乗分布に従う。

一方、 S_B は、

$$\begin{aligned} S_B &= \sum_{i=1}^l n_i (\bar{y}_i - \bar{y})^2 = \sum_{i=1}^l n_i \{(\bar{y}_i - \mu) - (\bar{y} - \mu)\}^2 \\ &= \sum_{i=1}^l n_i \{(\bar{y}_i - \mu)^2 - 2(\bar{y}_i - \mu)(\bar{y} - \mu) + (\bar{y} - \mu)^2\} \\ &= \sum_{i=1}^l n_i (\bar{y}_i - \mu)^2 - n(\bar{y} - \mu)^2 \end{aligned}$$

と分解できるので、 $\frac{S_B}{\sigma^2}$ は、

$$\begin{aligned} \frac{S_B}{\sigma^2} &= \frac{1}{\sigma^2} \sum_{i=1}^l n_i (\bar{y}_i - \bar{y})^2 \\ &= \frac{1}{\sigma^2} \sum_{i=1}^l n_i (\bar{y}_i - \mu)^2 - \frac{n}{\sigma^2} (\bar{y} - \mu)^2 \\ &= \sum_{i=1}^l \left(\frac{\bar{y}_i - \mu}{\sqrt{\frac{\sigma^2}{n_i}}} \right)^2 - \left(\frac{\bar{y} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \right)^2 \end{aligned}$$

と示される。これより、

$$\sum_{i=1}^l \left(\frac{\bar{y}_i - \mu}{\sqrt{\frac{\sigma^2}{n_i}}} \right)^2 = \frac{S_B}{\sigma^2} + \left(\frac{\bar{y} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \right)^2$$

と示される。このとき、 $\alpha_i = \mu_i - \mu = 0$ のもとでは、 $\bar{y}_i$ は、平均 μ 、分散 $\frac{\sigma^2}{n_i}$ の正規分布に従い、 $\bar{y}$ は、平均 μ 、分散 $\frac{\sigma^2}{n}$

の正規分布に従うので、 $\frac{\bar{y}_i - \mu}{\sqrt{\frac{\sigma^2}{n_i}}}$ 、および、 $\frac{\bar{y} - \mu}{\sqrt{\frac{\sigma^2}{n}}}$ は、ともに標準正規分布に従う。このため、 $\sum_{i=1}^l \left(\frac{\bar{y}_i - \mu}{\sqrt{\frac{\sigma^2}{n_i}}} \right)^2$ は、自由

度 l のカイ2乗分布に従い、 $\left(\frac{\bar{y} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \right)^2$ は、自由度1のカイ2乗分布に従う。さらに、 S_B と $\bar{y}$ は確率的に独立なことが示さ

れるので、カイ2乗分布の再生性より、 $\frac{S_B}{\sigma^2}$ は、自由度 $l-1$ のカイ2乗分布に従う。

この解答速報の著作権はT A C (株)のものであり，無断転載・転用を禁じます。

- (3) (2) より， S_W/σ^2 は自由度 $n-l$ のカイ2乗分布に従い， $\alpha_i = \mu_i - \mu = 0$ のもとでは， S_B/σ^2 は自由度 $l-1$ のカイ2乗分布に従う。このため，

$$F = \frac{\frac{\frac{S_B}{\sigma^2}}{n-l}}{\frac{\frac{S_W}{\sigma^2}}{l-1}} = \frac{\frac{S_B}{S_W}}{\frac{l-1}{n-l}}$$

は自由度 $(n-l, l-1)$ の F 分布に従う。