統計検定　2級　解答解説（2018年11月実施：第13回）

問 1（解答番号 $\boxed{1}$〜$\boxed{3}$ ）

[1]

（ア）: $100 - 85.1 - 2.1 = 12.8$ です。

（イ）: $100 - 76.6 - 17.0 - 2.1 = 4.3$ です。

$\boxed{1}$の正解は⑤となります。

[2]

C : 120〜140 にデータがあるので, 2017 年に相当します。

B : 100 以上のところにデータがあるので, 1985 年に相当します。

以上より, $\boxed{2}$の正解は①となります。

[3]

I : 四分位範囲は, 箱の長さに相当します。I は誤りです。

II : 1952 年（A）の最大値は 70 以上で, 1985 年（C）の最大値は 140 以下ですので, 誤りです。

III : 箱ひげ図における, 箱の中央の線が中央値です。正しいです。

以上より, $\boxed{3}$の正解は③となります。

問 2（解答番号 $\boxed{4}$ ）

I : 正しいです。

II : 20〜54 歳ではそれ以降の年齢に比べ直線関係がみてとれますので, 相関係数はむしろ大きくなると考えられます。誤りです。

III : 回帰係数よりわかります。誤りです。

$\boxed{4}$の正解は①となります。

問 3（解答番号 $\boxed{5}$〜$\boxed{6}$ ）

[1]

求める値を x と表しますと,

$$\frac{111.7 - x}{x} \times 100 = 4.98$$

より, $x \approx 106$ と求まります。$\boxed{5}$の正解は①となります。

[2]

10 月とその前後（9 月と 11 月）のデータの平均をとります。$\boxed{6}$の正解は④となります。

問 4（解答番号 $\boxed{7}$）

ラスパイレス価格指数は, 基準年の数量を用いて計算されます。$\boxed{7}$の正解は②となります。

問 5（解答番号 $\boxed{8}$）

I : N 個の個体から成る有限母集団から n 個の標本を非復元無作為抽出するとき, ある個体が標本の中にある確率は, $\dfrac{{}_{N-1}C_{n-1}}{{}_{N}C_{n}}$ でどの個体でも同じ, またある n 個の個体の組が抽出される確率は常に $1/{}_{N}C_{n}$ で同じになります。正しいです。

II : 比例抽出法の場合などに限ります（標本サイズにかかわらず, が誤り）。誤りです。

III : ある層だけ極端に個体数が少ない場合などが考えられます。正しいです。

$\boxed{8}$の正解は③となります。

問 6（解答番号 $\boxed{9}$）

二段抽出法のことです。$\boxed{9}$の正解は②となります。

問 7（解答番号 $\boxed{10}$〜$\boxed{11}$）

[1]

求める確率は, $0.7 \times 0.02 + 0.3 \times 0.08 = 0.038$ となります。$\boxed{10}$の正解は②となります。

[2] クッキーが工場 A で製造されるという事象を A, カモノハシのクッキーが入っている事象を C と表しますと,

$$P_C(A) = \frac{P(A \cap C)}{P(C)} = \frac{0.7 \times 0.02}{0.038} \approx 0.368$$

となります。$\boxed{11}$の正解は②となります。

問 8（解答番号 $\boxed{12}$〜$\boxed{13}$）

[1]

Y は平均 $0.3 + 2x$, 分散 1 の正規分布に従います。

$$P(Y \geq 0) = P(0.3 + 2x + U \geq 0) = P(U \geq -2x - 0.3)$$

この確率が 0.95 です。標準正規分布の上側 95 %点が -1.645 なので, $-2x - 0.3 = -1.645$ を解い

て, $x \approx 0.67$ となります。$\boxed{12}$の正解は④となります。

[2]

Y の上側 5 %点を y と表しますと, $P(Y \geq y) = P(0.3+2x+U \geq y) = P(U \geq y-2x-0.3)$ となり, この確率が 0.05 なので, 今度は標準正規分布の上側 5 %点 1.645 を用いて, $y-2x-0.3 = 1.645$, つまり $y = 2x + 1.945$ となります。x と y が直線関係にあるグラフを選びます。

$\boxed{13}$の正解は①となります。

問 9（解答番号 $\boxed{14}$〜$\boxed{15}$ ）

[1]

$$\frac{P(X=x+1)}{P(X=x)} = \frac{{}_7C_{x+1}(2/3)^{6-x}(1/3)^{x+1}}{{}_7C_x(2/3)^{7-x}(1/3)^x} = \frac{\dfrac{7!}{(x+1)!(6-x)!}\cdot\dfrac{1}{3}}{\dfrac{7!}{x!(7-x)!}\cdot\dfrac{2}{3}} = \frac{-x+7}{2(x+1)}$$

より, $a = 7$, $b = 2$ となります。$\boxed{14}$の正解は②となります。

[2]

$\dfrac{P(X=x+1)}{P(X=x)} > 1$ を解くと $3x < 5$ となりますので, $x = 0, 1$ のとき, $P(X = x+1) > P(X = x)$ です。つまり, $P(X = 0) < P(X = 1) < P(X = 2)$ です。

同様に, $\dfrac{P(X=x+1)}{P(X=x)} < 1$ より $x \geq 2$ となります。$x \geq 2$ のとき, $P(X = x+1) < P(X = x)$ です。つまり, $P(X = 2) > P(X = 3) > \cdots$ となります。

以上より, $x = 2$ で $P(X = x)$ は最大になります。$\boxed{15}$の正解は②となります。

問 10（解答番号 $\boxed{16}$ ）

$$E[\overline{X}] = \frac{1}{n}(E[X_1] + E[X_2] + \cdots + E[X_n]) = \frac{1}{n}(\mu + \mu + \cdots + \mu) = \mu$$

です。また, 各 X は互いに独立なので,

$$V[\overline{X}] = \frac{1}{n^2}(V[X_1] + V[X_2] + \cdots + V[X_n]) = \frac{1}{n^2}(\sigma^2 + \sigma^2 + \cdots + \sigma^2) = \frac{\sigma^2}{n}$$

となります。$\boxed{16}$の正解は③となります。

問 11（解答番号 $\boxed{17}$～$\boxed{19}$ ）

[1]

正規分布の歪度は 0, 尖度は, 問題文の定義によりますと, 0 です（正規分布の μ_4/σ^4 の値は常に 3 です)。$\boxed{17}$の正解は①となります。

[2]

一様分布は平均に関して対称ですので, 歪度は 0 になります。尖度は, 正規分布よりも中心部が平坦で, 裾が短いため, 負の値をとると考えられます。念のため計算しますと, 平均が 0 であることから,

$$\sigma^2 = E[X^2] = \int_{-1}^{1} x^2 \cdot \frac{1}{2} dx = \frac{1}{3}$$

$$E[(X-\mu)^4] = E[X^4] = \int_{-1}^{1} x^4 \cdot \frac{1}{2} dx = \frac{1}{5}$$

より, 尖度は $(1/5)/(1/3^2) - 3 = -1.2$ となります。$\boxed{18}$の正解は⑤となります。

[3]

I : 逆です。右に裾が長いと, 平均より右側の (平均より大きい) データの影響が大きくなり, $(X-\mu)^3$ の値は正に大きくなっていく, と考えるとよいです。

II : 逆です。裾が長いほど, 絶対値の大きいデータの影響を受け, 大きくなります。

III : t 分布は平均に関して対称なので, 歪度は 0, ここまでは正しいです。自由度が大きくなるほど, 標準正規分布（尖度 0）に近づいていきますので, 尖度の絶対値は小さくなります。

$\boxed{19}$の正解は⑤となります。

問 12（解答番号 $\boxed{20}$ ）

母比率の区間推定です。母比率を p, 標本比率を $\hat{p}$, 標本数を n とすると, n が十分大きいとき, p の 95 %信頼区間は,

$$\hat{p} - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

です。$\hat{p} = 0.02$, $n = 1338$ を代入します。$\boxed{20}$の正解は②となります。

問 13（解答番号 $\boxed{21}$ ）

（ア）検定統計量は, $t = \dfrac{\overline{X} - \mu}{\sqrt{U^2/n}}$ となります (μ : 母平均, $\overline{X}$: 標本平均, U^2 : 不偏分散, n : 標本数)。

（イ）自由度 19 の t 分布を用います。対立仮説より, 両側検定になりますので, 上側 2.5 %点 2.093, 上側 97.5 %点 -2.093 を用います。

（ウ）（ア）の t 値を計算するとおよそ -1.78 となり, 上側 97.5 %点と上側 2.5 %点の間にあるので, 帰無仮説を棄却しません。

$\boxed{21}$ の正解は④となります。

問 14（解答番号 $\boxed{22}$ ～ $\boxed{23}$ ）

[1]

母分散の比の検定です。第 1 自由度は $30-1$, 第 2 自由度は $31-1$ です。$\boxed{22}$ の正解は⑤となります。

[2]

第 1 種の過誤の確率とは, 帰無仮説 H_0 が正しいのに, H_0 を棄却する確率です。その確率ですが, H_0 が正しいとき,

$$\begin{aligned}
&(3\text{ つの検定のどれか 1 つでも } H_0\text{を棄却する確率})\\
&= 1-(3\text{ つの検定のいずれも } H_0\text{を棄却しない確率})\\
&= 1-(1\text{ つの組み合わせの検定で } H_0\text{を棄却しない確率})^3\\
&= 1-(1-0.05)^3 \approx 0.14
\end{aligned}$$

となります。$\boxed{23}$ の正解は④となります。

問 15（解答番号 $\boxed{24}$ ～ $\boxed{26}$ ）

[1]

X は二項分布 $B(200, 0.05)$ に従います。その期待値は $200 \times 0.05 = 10$, 分散は $200 \times 0.05 \times (1-0.05) = 9.5$ となります。$\boxed{24}$ の正解は③となります。

[2]

不良品の母比率を r, 標本比率を $\hat{r}$ とすると, $\hat{r}$ は, 標本数 n が十分大きいとき, 平均 r, 分散 $r(1-r)/n$ の正規分布に従うと近似できます。帰無仮説 H_0 のもとで, $r = 0.05$, $n = 200$ を代入しますと, その平均は 0.05, 分散は 0.0002375 となります。

さて, 対立仮説より右片側検定を行いますので, 片側 P 値は,

$$P(\hat{r} > 16/200) = P\left(\frac{\hat{r}-0.05}{\sqrt{0.0002375}} > \frac{16/200-0.05}{\sqrt{0.0002375}}\right) \approx P(Z > 1.95) \approx 0.26$$

となります（Z は標準正規分布に従う確率変数)。$\boxed{25}$ の正解は②となります。

[3]

それぞれのメーカーの標本比率は, $\hat{r}_A = 16/200$, $\hat{r}_B = 17/200$ です。よって, $\hat{d} = \hat{r}_A - \hat{r}_B = -0.005$ となります。

ところで, $\hat{r}_A$ は正規分布 $N\left(r_A, \dfrac{r_A(1-r_A)}{n_A}\right)$ に, $\hat{r}_B$ は正規分布 $N\left(r_B, \dfrac{r_B(1-r_B)}{n_B}\right)$ に従うので, $\hat{d}$ は正規分布 $N\left(r_A - r_B, \dfrac{r_A(1-r_A)}{n_A} + \dfrac{r_B(1-r_B)}{n_B}\right)$ に従います (なお, $n_A = n_B = 200$)。帰無仮説 H_0 :「$d = r_A - r_B = 0$」のもとで, 分散の中で $r_A = r_B = \dfrac{16+17}{200+200} = 0.0825$ と近似すると, $\hat{d}$ は, 正規分布 $N(0, 0.000825)$ に従います。

対立仮説より両側検定なので, 求める両側 P 値は,

$$2 \times P(\hat{d} < -0.005) = 2 \times P\left(Z < -\frac{0.005}{\sqrt{0.000825}}\right) \approx 2 \times P(Z < -0.174) \approx 2 \times 0.43 = 0.86$$

と求まります。$\boxed{26}$ の正解は⑤となります。

問 16(解答番号 $\boxed{27}$ ~ $\boxed{28}$)

[1]

理論度数は, どのセルも $102/6 = 17$ です。χ^2 統計量は $\sum \dfrac{(観測度数 - 理論度数)^2}{理論度数}$ です。$\boxed{27}$ の正解は①となります。

[2]

自由度は, (セルの数) $- 1 = 6 - 1 = 5$ です。また, 右片側検定を行いますので, 自由度 5 の χ^2 分布の上側 5 %点 11.07 と比較します。約 2.59 はこの値より小さいので, 帰無仮説を棄却しません。

$\boxed{28}$ の正解は③となります。

問 17(解答番号 $\boxed{29}$ ~ $\boxed{31}$)

[1]

F 値の第 2 自由度が, 52 ですが, これは (標本数) $-$ (説明変数の数) $- 1$ にあたります。いま説明変数の数は 2 なので, 標本数は 55 となります。$\boxed{29}$ の正解は④となります。

[2]

I : 誤りです。「Intercept」の「Std.error」が, α の標準誤差です。1.137×10^2 です。

II : 正しいです。P 値が 0.05 より小さければ有意 (0 でない) とみなします。いま, どの値も 0.05 より小さいです。

III : 誤りです。「Adjusted R-squared」の値が, 自由度調整済み決定係数です。

$\boxed{30}$ の正解は②となります。

[3]

I : 正しいです。β_1 の推定値が -6.617×10^{-2} で負です。

II : β_2 の推定値が正であること, GDP が増えれば logGDP も増えることから, 正しいです。

III：回帰式に推定値を代入します。

$$-1.283 \times 10^3 - -6.617 \times 10^{-2} \times 400 + 1.757 \times 10^2 \times 10 = 447.532$$

となります。正しいです。

$\boxed{31}$の正解は④となります。

問 18（解答番号 $\boxed{32}$〜$\boxed{34}$ ）

[1]

I：正しいです。残差平方和を S としますと，

$$\hat{\sigma}_u^2 = \frac{S}{n-k-1}$$

という関係があります。n は標本数でいまは 5, k は説明変数の数でいまは 1 です。よって，$S = (5-1-1) \times 0.608^2 \approx 1.1$ となります。

II：誤りです。t 値は変わりません。

III：正しいです。

$\boxed{32}$の正解は④となります。

[2]

I：誤りです。z の P 値は，切片や x の P 値と比べて小さいです。

II：誤りです。説明変数間の相関が高いと，多重共線性の問題が生じます。

III：誤りです。P 値の方が有意水準より大きいため，帰無仮説を棄却しません。

$\boxed{33}$の正解は⑤となります。

[3]

I：誤りです。同じ値になるとは限りません。

II：正しいです。

III：誤りです。b' が有意でないので，「x が変化しても y は変化しない」と解釈されます。

$\boxed{34}$の正解は①となります。

（以上です）