統計検定２級　全 10 回のオリジナル解説より抜粋

【第１回：2011 年 11 月実施】

問 5（解答番号 $\boxed{9}$ ）

30 匹の小型犬の体重を $x_1, x_2, \cdots, x_{30}$, それらの相加平均を $\overline{x}$, 標準偏差を s と表しますと,

$$\overline{x} = \frac{1}{30}\sum_{i=1}^{30} x_i = 4.5$$
$$s^2 = \frac{1}{30}\sum_{i=1}^{30}(x_i - \overline{x})^2 = 1.0^2$$

であり, 変動係数 $a = s/\overline{x} = 1.0/4.5$ となります。ここで, 各犬の体重が 0.5kg ずつ増加したとして,

それら新たな体重を $x'_1, x'_2, \cdots, x'_{30}$ と表しますと, $x'_i = x_i + 0.5$ であり, 新たな相加平均を $\overline{x}'$, 標準偏差を s' と表しますと, $\overline{x}' = \overline{x} + 0.5$, $s = s'$ が成り立ちます。よって,

$$(\text{新たな変動係数}) = s'/\overline{x}' = s/(\overline{x} + 0.5) = 1.0/5.0 = \frac{5.0}{4.5}a = \frac{10}{9}a$$

となります。

よって, $\boxed{9}$ の正解は 3 となります。

問 14（解答番号 $\boxed{18}$ ）

平均 μ, 分散 σ^2 の正規分布の確率密度関数 $f(x)$ は,

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

です。

a. $x = \mu$ において $f(x)$ は最大値 $\dfrac{1}{\sqrt{2\pi}\sigma}$ をとります。

b. σ が大きいほど, その最大値は大きいです。よって, a. と b. は正しいです。

c. X を $N(0, \sigma_1^2)$ に従う確率変数, Y を $N(0, \sigma_2^2)$ に従う確率変数, Z を標準正規分布に従う確率変数としますと, $Pr(X \geq x) = Pr(Z \geq x/\sigma_1)$, $Pr(Y \geq x) = Pr(Z \geq x/\sigma_2)$ と書き直せます。$x \geq 0$, $\sigma_1 < \sigma_2$ から $x/\sigma_1 \geq x/\sigma_2$ なので,

$Pr(Z \geq x/\sigma_1) \leq Pr(Z \geq x/\sigma_2)$, すなわち $Pr(X \geq x) \leq Pr(Y \geq x)$ です。よって c. も正しいです。

以上より，$\boxed{18}$ の正解は 5 となります。

問 17（解答番号 $\boxed{21}$ ）

母支持率を p, 標本支持率を $\hat{p}$, 標本数を n と表しますと, 通常 n は非常に大きいので, 中心極限定理を用いて,

$\hat{p}$ は正規分布 $N\left(p, \dfrac{p(1-p)}{n}\right)$ に従うと近似できます。よって, 例えば p の 95 %信頼区間は,

$$Pr\left(p - Z_{0.025}\sqrt{\frac{p(1-p)}{n}} < \hat{p} < p + Z_{0.025}\sqrt{\frac{p(1-p)}{n}}\right) = 0.95$$

の () の中を, 分散に現れる p を $\hat{p}$ で近似し, そののち p について整理した,

$$\hat{p} - Z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + Z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

となります。この信頼区間の幅は, $2Z_{0.025}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$ です。n が 4 倍になれば, この幅は半分になることがわかります。信頼係数が変わっても一緒です。

以上より，$\boxed{21}$ の正解は 4 となります。

【第 2 回：2012 年 11 月実施】

問 13（解答番号 $\boxed{20}$ ）

n 個の世帯について, 各世帯の人数を $X_1, X_2, \cdots, X_n$ と表しますと, 「世帯毎の人数」の推定量は $\overline{X} = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$ となります。すると, 人口の推定量は $100000\overline{X}$ となります。その変動係数は, $E[\overline{X}] = E[X_i]$, $V[\overline{X}] = \dfrac{V[X_i]}{n}$ などより,

$$\frac{\sqrt{V[100000\overline{X}]}}{E[100000\overline{X}]} = \frac{1}{\sqrt{n}} \frac{\sqrt{V[X_i]}}{E[X_i]}$$

となります。

ところで, 母集団の各世帯における人数の母平均を μ, 母分散を σ^2 と表しますと,

$$E[X_i] = \mu, \quad V[X_i] = \sigma^2$$

ですので, 上記の変動係数は

$$\frac{1}{\sqrt{n}} \underline{\frac{\sigma}{\mu}}$$

となりますが, ここで母集団の変動係数（下線部）が 1.0 以下なので,

$$\frac{1}{\sqrt{n}} \leq 0.05$$

より $n \geq 400$ となります。$\boxed{20}$ の正解は②となります。

問 14（解答番号 $\boxed{21}$～$\boxed{22}$ ）

[1]

入場者数を N 人としますと, i 番目の人の通し番号 X_i は区間 $(1, N)$ の離散型一様分布に従います。その期待値は $E[X_i] = \dfrac{N+1}{2}$ です。よって, 通し番号の合計 $S = X_1 + X_2 + \cdots + X_{300}$ の期待値は $E[S] = 300 \times \dfrac{N+1}{2}$ です。いま, N の推定量を $\hat{N}$ と表し,

$$\hat{N} = 2 \times \frac{S}{300} - 1$$

としますと, $E[\hat{N}] = N$ となり, N の不偏推定量になります。よって, $2 \times \dfrac{4510384}{300} - 1 \approx 30000$ が答えとなります。$\boxed{21}$の正解は③となります。

[2]

N の標準誤差とは, $\hat{N}$ の標準偏差です。

$$V[\hat{N}] = \frac{2^2}{300^2} \times 300V[X_i]$$

ですが, ここで $V[X_i]$ は,

$$\begin{aligned} E[X_i^2] &= \frac{1}{N}(1^2 + 2^2 + \cdots + N^2) \\ &= \frac{1}{N} \times \frac{1}{6}N(N+1)(2N+1) \\ &= \frac{1}{6}(N+1)(2N+1) \end{aligned}$$

より,

$$\begin{aligned} V[X_i] &= E[X_i^2] - (E[X_i])^2 \\ &= \frac{1}{6}(N+1)(2N+1) - \frac{1}{4}(N+1)^2 \\ &= \frac{1}{12}(N-1)(N+1) \end{aligned}$$

となります。$N \approx 30000$ を代入して, $\sqrt{V[\hat{N}]} \approx 1000$ を得ます。$\boxed{22}$の正解は②となります。

なお, N が非常に大きいことから, 離散型の一様分布ではなく連続型の一様分布とみなし,

$$V[X_i] = \frac{(N-1)^2}{12}$$

として, $N \approx 30000$ を代入しても同じ答えが得られます。

【第3回：2013年11月実施】

問4（解答番号 $\boxed{5}$ ）

相関係数は, 各変数を標準化しても値は変わりませんので, A と C の値は同じになります。なお, D は, x, y を標準化したものをそれぞれ x', y' とおくと,

$$x_i' = \frac{x_i - \overline{x}}{s_x}, \quad y_i' = \frac{y_i - \overline{y}}{s_y}$$

であり, $\overline{x}' = \overline{y}' = 0$ です。よって,

$$\begin{aligned} Cov(x', y') &= \frac{1}{n-1} \sum (x_i' - \overline{x}')(y_i' - \overline{y}') \\ &= \frac{1}{n-1} \sum x_i' \cdot y_i' \\ &= \frac{1}{n-1} \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{s_x \cdot s_y} \end{aligned}$$

となりますが, 最後の式に問題文にある s_x と s_y を代入すると, r_{xy} と等しくなることがわかります。

以上より, $\boxed{5}$ の正解は⑤となります。

なお, 本問では分散, 共分散ともに不偏推定量を用いています。不偏でない分散・共分散（$n-1$ で割らずに n で割るもの）か, それとも本問のような分散・共分散か, テキストや試験によってどちらを使うかマチマチですので, 問題文をよく読んで間違えないようにして下さい（統計検定試験は不偏推定量で統一されています)。

問16（解答番号 $\boxed{25}$～$\boxed{26}$ ）

[1]

標本比率 $\hat{p}$, 標本数 n と表すと, 母比率 p の 95 %信頼区間は,

$$\hat{p} \pm 1.96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

で表されます。$\hat{p} = 0.25$, $n = 600$ なので, $\boxed{25}$ の正解は⑤となります。

[2] 信頼区間の幅は,

$$2 \times 1.96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

ですので,

$$2 \times 1.96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq 0.04$$

すなわち

$$n \geq \left(\frac{2 \times 1.96}{0.04}\right)^2 \times \hat{p}(1-\hat{p})$$

となります。$\hat{p}(1-\hat{p})$ は $0 \leq \hat{p} \leq 0.3$ において単調増加なので, $\hat{p} = 0.3$ のときに最大になりますので,

$$n \geq \left(\frac{2 \times 1.96}{0.04}\right)^2 \times (0.3 \times 0.7) = 2016$$

でなければなりません。$\boxed{26}$ の正解は④となります。

【第 4 回：2014 年 6 月実施】

問 13（解答番号 $\boxed{20}$〜$\boxed{22}$）

[1]

3 の目が出る回数 r は二項分布 $B(7, 1/6)$ に従い,

$$ {}_7C_4 \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)^3 $$

が求める確率になります。${}_7C_4 = 35$ なので, $\boxed{20}$ の正解は④となります。

[2]

帰無仮説が正しいとしたときに,「3 の目が 4 回出る」という事象が起こる確率が非常に小さければ, 滅多に起こらないことが起こったということで, 帰無仮説を疑うことになります。このとき, 5 回以上の確率も加えた確率を P 値とし, 有意水準である 5 %と比較します。

$\boxed{21}$ の正解は①となります。

[3]

1 の目が 4 回出る確率も [1] と同じ式で計算できて 0.015629, 2 の目が 4 回出る確率も同様に 0.015629, $\cdots$, です。これらの事象は互いに排反なので, 求める確率は, $0.015629 \times 6 \approx 0.094$ となります。

$\boxed{22}$ の正解は③となります。

問 18（解答番号 $\boxed{32}$〜$\boxed{33}$）

[1]

標本回帰方程式によると, 気温が 1 度上がると傾きの大きさだけ（4.8617）降水量が増えます。$\boxed{32}$ の正解は④となります。

[2]

決定係数は B 君の分析の方が大きいので, I は誤りです。

切片や平均気温の標準誤差をみると, B 君の分析の方が小さくなっています。これらは, 標本回帰方程式における切片と傾きの標準偏差です。II は正しいです。

自由度修正済み決定係数は, 説明変数の数の検討に用いられます。単回帰分析同士のこの決定係数の比較は意味がありません。III は誤りです。

以上より, $\boxed{33}$ の正解は②となります。

【第 5 回：2014 年 11 月実施】

問 6（解答番号 $\boxed{8}$〜 $\boxed{11}$ ）

[1]

以下のデータの平均を求めることになります。

$$\underbrace{0,0,0,\cdots,0}_{55\text{ 個}},\underbrace{1,1,\cdots,1}_{144\text{ 個}},\underbrace{2,2,\cdots,2}_{140\text{ 個}},\cdots,\underbrace{6,6,6,6,6,6}_{6\text{ 個}}$$

よって,

$$(0\times 55+1\times 144+2\times 140+\cdots+6\times 6)\div 500=2.00$$

となります。$\boxed{8}$の正解は③です。

[2]

母数 λ のポアソン分布の確率関数は,

$$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$$

です。$\boxed{9}$の正解は⑤です。

[3]

母数 λ のポアソン分布の期待値 $E[X]$, 分散 $V[X]$ はともに λ です。ここで $V[X]=E[X^2]-(E[X])^2$ ですので,

$$E[X^2]=V[X]+(E[X])^2=\lambda+\lambda^2$$

になります。$\boxed{10}$の正解は④です。

[4]

検定統計量としては,

$$\chi^2=\frac{(55-67.7)^2}{67.7}+\frac{(144-135.3)^2}{135.3}+\cdots+\frac{(6-8.3)^2}{8.3}\approx 4.498$$

を用い, 自由度 5 の χ^2 分布の上側%点と比較します。

自由度に関して, セルが 7 つあるので, $7-1=6$ とし, さらに, 母数を 1 つ最尤推定値で推定しているので, 1 減らして, 5 とします（推定した母数の数だけ自由度を減じます)。

以上より, $\boxed{11}$の正解は⑤です。

問 12（解答番号 $\boxed{32}$〜 $\boxed{35}$ ）

[1]

　やや直感に頼るところもあります。$\boxed{32}$の正解は④になります。

[2-1]

　決定係数が該当します。$\boxed{33}$の正解は①になります。

[2-2]

　母回帰方程式

$$Y = \alpha + \beta X + \varepsilon$$

において, 誤差項 ε が標本毎に独立に正規分布 $N(0, \sigma^2)$（σ^2 は定数）に従うとしますと, 回帰係数の推定量（標本から得られる回帰係数）$\hat{\alpha}$, $\hat{\beta}$ は正規分布に従います。

　本問の場合, 詳細は非常に長くなりますので割愛しますが, 男子, 女子それぞれの回帰係数の推定量 $\hat{\beta}^{(1)}$, $\hat{\beta}^{(2)}$ は正規分布に従います。すると, $\hat{\beta}^{(1)} - \hat{\beta}^{(2)}$ も正規分布に従います。このことを利用して, それぞれの母回帰係数について「$\beta^{(1)} = \beta^{(2)}$」を検定することができます（実際は t 分布を用います)。

　$\boxed{34}$の正解は②になります。

[3]

　本問における重回帰モデルは, 体重 Y, 身長 X, 性別 Z（男性ならば $Z = 1$, 女性ならば $Z = 0$）と表しますと,

$$Y = \alpha + \beta^{(1)} X + \beta^{(2)} Z + \beta^{(3)} XZ + \varepsilon \cdots (a)$$

となっています。

　I は, 誤りです。本問のように, ダミー変数として用いて, 意味のある分析ができます。

　II について, (a) の式で $Z = 0$ としますと, 回帰直線 $Y = \alpha + \beta^{(1)} X$ が現れます。II は誤りです。偶然ではありません。

　III について, 男性のデータを用いる, つまり (a) で $Z = 1$ としますと, 回帰直線 $Y = \alpha + \beta^{(2)} + (\beta^{(1)} + \beta^{(3)}) X$ が現れます。女性のデータを用いる, つまり (a) で $Z = 0$ としますと, 回帰直線 $Y = \alpha + \beta^{(1)} X$ が現れます。

　$\beta^{(3)}$ は, 男性と女性の回帰係数の差に対応しているといえます。III は正しいです。

　以上より, $\boxed{35}$の正解は③になります。

【第 6 回：2015 年 6 月実施】

問 8（解答番号 $\boxed{16}$ 〜 $\boxed{18}$ ）

[1]

$$\frac{5 \times 7 \times 3}{{}_{15}C_3} = \frac{3}{13}$$

となり, $\boxed{16}$ の正解は⑤となります。

[2]

100 円玉が 2 枚と, あと 1 枚（何でもよい）あればよいです。100 円玉 3 枚をとる場合, 100 円玉 2 枚と 10 円玉 1 枚をとる場合, 100 円玉 2 枚と 1 円玉 1 枚をとる場合をそれぞれ考えて,

$$\frac{{}_5C_3 + {}_5C_2 \times 7 + {}_5C_2 \times 3}{{}_{15}C_3} = \frac{22}{91}$$

となります。$\boxed{17}$ の正解は④となります。

[3]

取り出した 3 枚の合計が 150 円以上で, かつ 1 円玉が含まれている確率は, $\dfrac{{}_5C_2 \times 3}{{}_{15}C_3}$ となります。この確率を [2] の確率で割ると, $\dfrac{3}{11}$ となり, これが答えです。$\boxed{18}$ の正解は②となります。

問 11（解答番号 $\boxed{23}$ ）

二項分布を用いて解くと計算が大変なことになりますので, $n = 36$ と試行回数が小さいのですが, 正規近似をします。正解した数を S と表しますと, S は二項分布 $B(36, 1/2)$ に従い, ここで正規近似を行いますと, S は期待値 $36 \times \dfrac{1}{2} = 18$, 分散 $36 \times \dfrac{1}{2} \times \left(1 - \dfrac{1}{2}\right) = 9$ の正規分布に従います。よって,

$$P(S \geq 24) = P\left(\frac{S - 18}{3} \geq \frac{24 - 18}{3}\right) = P(Z \geq 2) = 0.0228$$

となります（Z は標準正規分布に従う確率変数)。

ここで, 試行回数が小さいので, 連続補正を行いますと,

$$P(S \geq 24 - 0.5) = P(Z \geq 1.83) = 0.0336$$

となります。なお, 実際に二項分布のまま計算もしてみましたが, 約 3 %になりました（試験時間内で計算するのは難しい)。$\boxed{23}$ の正解は⑤となります。

問 13（解答番号 $\boxed{25}$〜 $\boxed{26}$ ）

[1]

母分散未知の場合の母平均の検定です。$\boxed{25}$の正解は③となります。

[2]

第二種の過誤率とは, H_1 が正しいときに H_0 を採択してしまう確率です。いま, H_1 が何であれ, H_0 を採択してしまう確率は, 一様乱数の値が 0.01 より大きい確率なので, 0.99 になります。

なお, 第一種の過誤率とは, H_0 が正しいときに H_0 を棄却する確率であり, 有意水準に一致しますので, いまは 0.01 になり, 不合理な値ではありません。

$\boxed{26}$の正解は③となります。

【第 7 回：2015 年 11 月実施】

問 11（解答番号 $\boxed{21}$〜$\boxed{22}$）

[1]

1000 歩中のカウントされた歩数を S と表しますと, S は二項分布 $B(1000, 1/10)$ に従います。二項近似を行いますと, S は正規分布 $N\left(1000 \times \frac{1}{10}, 1000 \times \frac{1}{10} \times \frac{9}{10}\right)$, つまり $N(100, 90)$ に従います。よって

$$P(S \geq 110) = P\left(\frac{S-100}{\sqrt{90}} \geq \frac{110-100}{\sqrt{90}}\right) \approx P(Z > 1.054) \approx 0.15$$

となります。$\boxed{21}$の正解は②となります。

[2]

「カウント数が 67 である」という事象を A,「600 歩あるく」という事象を B,「700 歩あるく」という事象を C としますと,

$$\begin{aligned} P(B \mid A) &= \frac{P(A \cap B)}{P(A)} = \frac{P(A \mid B)P(B)}{P(A \mid B)P(B) + P(A \mid C)P(C)} \\ &= \frac{\frac{1}{2} \times {}_{600}C_{67} \left(\frac{1}{10}\right)^{67} \left(\frac{9}{10}\right)^{600-67}}{\frac{1}{2} \times {}_{600}C_{67} \left(\frac{1}{10}\right)^{67} \left(\frac{9}{10}\right)^{600-67} + \frac{1}{2} \times {}_{700}C_{67} \left(\frac{1}{10}\right)^{67} \left(\frac{9}{10}\right)^{700-67}} \end{aligned}$$

となり, ${}_{600}C_{67} \left(\frac{1}{10}\right)^{67} \left(\frac{9}{10}\right)^{600-67} = 0.033$, ${}_{700}C_{67} \left(\frac{1}{10}\right)^{67} \left(\frac{9}{10}\right)^{700-67} = 0.048$ となるのですが, これを試験時間内に解くのは大変難しいです。

そこで, 正規近似を行います。

$$\begin{aligned} P_1 &= {}_{600}C_{67} \left(\frac{1}{10}\right)^{67} \left(\frac{9}{10}\right)^{600-67} \\ &\approx P(66.5 < S < 67.5) \qquad (S \text{ はこのとき二項分布 } B(600, 1/10) \text{ に従っている}) \\ &= P\left(\frac{66.5 - 600 \times 1/10}{\sqrt{600 \times 1/10 \times 9/10}} < Z < \frac{67.5 - 600 \times 1/10}{\sqrt{600 \times 1/10 \times 9/10}}\right) \\ &\approx P(0.88 < Z < 1.02) = 0.1894 - 0.1539 = 0.0355 \end{aligned}$$

$$\begin{aligned} P_2 &= {}_{700}C_{67} \left(\frac{1}{10}\right)^{67} \left(\frac{9}{10}\right)^{700-67} \\ &\approx P(66.5 < S < 67.5) \qquad (S \text{ はこのとき二項分布 } B(700, 1/10) \text{ に従っている}) \\ &= P\left(\frac{66.5 - 700 \times 1/10}{\sqrt{700 \times 1/10 \times 9/10}} < Z < \frac{67.5 - 700 \times 1/10}{\sqrt{700 \times 1/10 \times 9/10}}\right) \\ &\approx P(-0.44 < Z < -0.31) = 0.3783 - 0.33 = 0.0483 \end{aligned}$$

となります。よって, 求める答えは,

$$\frac{\frac{1}{2}\times 0.0355}{\frac{1}{2}\times 0.0355+\frac{1}{2}\times 0.0483}\approx 0.42$$

となります。$\boxed{22}$の正解は③となります。

問 13（解答番号 $\boxed{24}$ ）

I : 正しいです（以下の定義も参照）。

II : 正しいです。$E[X_1]=\mu$ ですが一致推定量ではありません。

III : まずは, 以下の一致性の定義などをご覧下さい。

一致性：母数 θ の推定量 $\hat{\theta}$ について, 標本数 n が大きくなるにつれ, 任意の正の数 ε（>0）に対して,

$$P\left(|\hat{\theta}-\theta|>\varepsilon\right)\to 0$$

が成り立つとき, $\hat{\theta}$ は母数 θ の一致推定量である, という。

なお, 標本平均 $\overline{X}$ は μ の一致推定量である。このことは, チェビシェフの不等式を用いて示すことができる。母分散を σ^2 と表すと, $E[\overline{X}]=\mu$, 分散 $V[\overline{X}]=\dfrac{\sigma^2}{n}$ であるから,

$$P\left(|\overline{X}-E[\overline{X}]|>k\sqrt{V[\overline{X}]}\right)<\frac{1}{k^2}\Longleftrightarrow P\left(|\overline{X}-\mu|>\frac{k\sigma}{\sqrt{n}}\right)<\frac{1}{k^2}$$

となるが, ここで $\dfrac{k\sigma}{\sqrt{n}}=\varepsilon$ とおくと,

$$P(|\overline{X}-\mu|>\varepsilon)<\frac{\sigma^2}{n\varepsilon^2}\to 0\quad(n\to\infty)$$

となり, 確かに $n\to\infty$ において右辺は 0 に収束し, 標本平均 $\overline{X}$ は μ の一致推定量になる。

さて, 例えば $\overline{X}'=\dfrac{3X_1-X_2+X_3+X_4+\cdots+X_n}{n}$ という推定量を考えますと, この期待値は μ になり, 分散は $\dfrac{n+8}{n^2}\sigma^2$ となります。上のやり方と同じように k を置くことで, 一致性をいうことができます。

したがって, III は誤りです。

$\boxed{24}$の正解は②となります。

問 17（解答番号 $\boxed{31}$ ）

第 1 種の過誤率は, H_0 が正しいときに H_0 を棄却する条件付確率です。

「H_0 が正しい」つまり「$p = 0.62$」

のときに,

「H_0 を棄却する」つまり「3 回とも針が上向きか, 3 回とも下向き（これらは互いに排反)」

となる確率なので, ②となります。

$\boxed{31}$の正解は②となります。

【第8回：2016年6月実施】

問3（解答番号 $\boxed{6}$〜$\boxed{8}$ ）

[1]

問題文を表に表しますと, 次のようになります。

人数	はい	いいえ	合計
属性 A			40
属性 B			60
合計	20	80	100

独立性の仮定のもとでは, 属性 B, いいえの欄の期待度数は $\dfrac{80 \times 60}{100} = 48$ となります。

$\boxed{6}$ の正解は④となります。

[2]

$(x, y) =$ (属性 A/B, はい/いいえ) としますと, $(0,0)$ が 40 個, $(0,1)$ が 10 個, $(1,0)$ が 10 個, $(1,1)$ が 40 個となります。x の平均を $\overline{x}$, などと表しますと,

$$\overline{x} = \frac{0 \times 40 + 0 \times 10 + 1 \times 10 + 1 \times 40}{100} = \frac{1}{2}$$
$$\overline{y} = \frac{0 \times 40 + 1 \times 10 + 0 \times 10 + 1 \times 40}{100} = \frac{1}{2}$$
$$\overline{xy} = \frac{0 \times 0 \times 40 + 0 \times 1 \times 10 + 1 \times 0 \times 10 + 1 \times 1 \times 40}{100} = \frac{2}{5}$$

より共分散は

$$\overline{xy} - \overline{x} \times \overline{y} = \frac{3}{20}$$

となります。次に,

$$\overline{(x^2)} = \frac{0^2 \times 40 + 0^2 \times 10 + 1^2 \times 10 + 1^2 \times 40}{100} = \frac{1}{2}$$
$$\overline{(y^2)} = \frac{0^2 \times 40 + 1^2 \times 10 + 0^2 \times 10 + 1^2 \times 40}{100} = \frac{1}{2}$$

より,x, y の分散は

$$V[x] = \overline{(x^2)} - (\overline{x})^2 = \frac{1}{4}$$
$$V[y] = \overline{(y^2)} - (\overline{y})^2 = \frac{1}{4}$$

となります。よって相関係数は,

$$\frac{3/20}{\sqrt{1/4}\sqrt{1/4}} = \frac{3}{5}$$

となります。

$\boxed{7}$の正解は①となります。

[3]

属性 B のデータをすべて 1/2 倍しても, 相関係数の値は変わりません。よってまず $r_1 = r_2$ です。

次に, r_3 について, 今度は $(1,0)$40 個, $(0,0)$10 個, $(1,1)$10 個, $(0,1)$40 個の相関係数を求めることになります。$\overline{x}$, $\overline{y}$, $V[x]$, $V[y]$ は r_1 のときと変わりませんが, $\overline{xy}$ のみ 1/10 となります（上記と同様に計算するとわかります)。すると, 共分散は $1/10 - (1/2)^2 = -3/20$ となり, r_1 と符号が逆になります。

以上より, $\boxed{8}$の正解は①となります。

問 7（解答番号 $\boxed{16}$〜$\boxed{18}$ ）

[1]

$$P(A \text{ が } B \text{ に勝つ}) \times P(A \text{ が } C \text{ に負ける}) \times P(A \text{ が優勝} \mid A \text{ が負ける}) = \frac{1}{2} \times \frac{1}{2} \times r = \frac{1}{4}r$$

です。

$\boxed{16}$の正解は⑤となります。

[2]

「[1] の場合の確率」: $\dfrac{1}{4}r$

「最初に A が 2 連勝する確率」: $\dfrac{1}{2} \times \dfrac{1}{2} = \dfrac{1}{4}$

「最初 B に負け, そのあと A が優勝する確率」: $\dfrac{1}{2} \times r$

これらは互いに排反なので, 足し合わせて, $P_A = \dfrac{1}{4} + \dfrac{3}{4}r$ となります。

$\boxed{17}$の正解は③となります。

[3]

A と B は対称的なので, $P_A = P_B$ です。また, C に関しては,

「最初に A が B に勝ち, 次に C が A に勝ち, 次に B にも勝ち優勝する」: $\dfrac{1}{8}$

「最初に B が A に勝ち, 次に C が B に勝ち, 次に A にも勝ち優勝する」: $\dfrac{1}{8}$

「最初に A が B に勝ち, 次に C が A に勝ち, 次に C が B に負け, その後優勝する」: $\dfrac{1}{8}r$

「最初に B が A に勝ち, 次に C が B に勝ち, 次に C が A に負け, その後優勝する」: $\dfrac{1}{8}r$

これらは互いに排反なので, 足し合わせて, $P_C = \dfrac{1}{4} + \dfrac{1}{4}r$ となります。これは P_A, P_B より小さいです。

$\boxed{18}$の正解は②となります。

問 8（解答番号 $\boxed{19}$〜$\boxed{20}$ ）

[1]

中央値は, $F(x) = 0.5$ のところですので, $x^2 = 0.5$ を解いて, 0.707 となります。

$\boxed{19}$の正解は④となります。

[2]

確率密度関数 $f(x)$ は, 分布関数 $F(x)$ を x で微分することで得られます。$0 \leq x \leq 1$ では, $f(x) = 2x$ で, それ以外では $f(x) = 0$ です。

$$E[X] = \int_0^1 x \times 2x dx = \left[\frac{2}{3}x^3\right]_0^1 = \frac{2}{3}$$
$$E[X^2] = \int_0^1 x^2 \times 2x dx = \left[\frac{1}{2}x^4\right]_0^1 = \frac{1}{2}$$

より,

$$V[X] = E[X^2] - (E[X])^2 = \frac{1}{18} \approx 0.056$$

となります。$\boxed{20}$の正解は②となります。

問 12（解答番号 $\boxed{29}$ ）

まず第 2 種の過誤率を求めます。第 2 種の過誤率は, 「H_1 が正しいときに, 誤って H_0 を採択する確率」です。

いま, 帰無仮説 H_0 :「$\mu = 600$」, 対立仮説 H_1 :「$\mu = 630$」とし, H_1 が正しいとします。

ところが検定を行う者はそんなことはわからず, 機械的に, $\dfrac{\overline{X} - 600}{10} < 2.33$ なら, H_0 を採択してしまいます。ここで, $\overline{X}$ は正規分布 $N(630, 50^2/25)$ に従いますので, この確率を求めると,

$$\begin{aligned}
&P\left(\frac{\overline{X} - 600}{10} < 2.33\right) \\
&= P(\overline{X} < 623.3) \\
&= P\left(\frac{\overline{X} - 630}{\sqrt{50^2/25}} < \frac{623.3 - 630}{\sqrt{50^2/25}}\right) \\
&= P(Z < -0.67) \quad (Z \text{ は標準正規分布に従う確率変数}) \\
&= 0.2514
\end{aligned}$$

となります。これが第 2 種の過誤率です。検出力 $= 1 -$ 第 2 種の過誤率 ですので, 検出力は約 0.75 となります。

$\boxed{29}$の正解は④となります。

【第 9 回：2016 年 11 月実施】

問 7（解答番号 $\boxed{12}$〜$\boxed{13}$ ）

[1]

$$
\begin{aligned}
&\underbrace{\frac{4}{10}}_{\text{通りかかった人が }A} \times \underbrace{\frac{40}{100}}_{A\text{ 型と予想する}} + \underbrace{\frac{3}{10}}_{\text{通りかかった人が }B} \times \underbrace{\frac{30}{100}}_{B\text{ 型と予想する}} + \underbrace{\frac{2}{10}}_{\text{通りかかった人が }O} \times \underbrace{\frac{20}{100}}_{O\text{ 型と予想する}} \\
&\quad + \underbrace{\frac{1}{10}}_{\text{通りかかった人が }AB} \times \underbrace{\frac{10}{100}}_{AB\text{ 型と予想する}} \\
&\quad = \frac{3}{10}
\end{aligned}
$$

となります。$\boxed{12}$の正解は②となります。

[2]

$$
\begin{aligned}
&P(\text{実際に }A\text{ 型である} \mid A\text{ 型であると予想する}) \\
&= \frac{P(\text{実際に }A\text{ 型である}, \text{かつ}, A\text{ 型であると予想する})}{P(A\text{ 型であると予想する})} \\
&= \frac{4/10 \times 2/3}{\underbrace{4/10}_{A\text{ 型である}} \times \underbrace{2/3}_{A\text{ 型の人を }A\text{ 型と予想}} + \underbrace{6/10}_{A\text{ 型でない}} \times \underbrace{1/3}_{A\text{ 型でない人を }A\text{ 型と予想}}} \\
&= \frac{4}{7} \approx 0.57
\end{aligned}
$$

となります。$\boxed{13}$の正解は③となります。

問 8（解答番号 $\boxed{14}$〜$\boxed{15}$ ）

[1]

$$
\begin{aligned}
Cov(X_1, Y) &= Cov\left(X_1, \frac{X_1 + X_2 + X_3}{3}\right) \\
&= \frac{1}{3} Cov(X_1, X_1 + X_2 + X_3) \\
&= \frac{1}{3}\{\underbrace{Cov(X_1, X_1)}_{=V[X_1]=1} + \underbrace{Cov(X_1, X_2)}_{=0} + \underbrace{Cov(X_1, X_3)}_{=0}\} \\
&= \frac{1}{3}
\end{aligned}
$$

であり, $V[X_1]=1$,

$$V[Y]=V\left[\frac{X_1+X_2+X_3}{3}\right]=\frac{1}{9}V[X_1+X_2+X_3]=\frac{1}{9}(V[X_1]+V[X_2]+V[X_3])=\frac{1}{3}$$

なので, 相関係数は,

$$\frac{Cov(X_1,Y)}{\sqrt{V[X_1]}\sqrt{V[Y]}}=\frac{1/3}{\sqrt{1/3}}=\frac{\sqrt{3}}{3}\approx 0.6$$

となります。$\boxed{14}$の正解は④となります。

[2]

$$\begin{aligned}
Cov(X_1,Y)&=Cov\left(X_1,\frac{X_1+X_2+X_3}{3}\right)\\
&=\frac{1}{3}Cov(X_1,X_1+X_2+X_3)\\
&=\frac{1}{3}\{\underbrace{Cov(X_1,X_1)}_{=V[X_1]=1}+\underbrace{Cov(X_1,X_2)}_{=0.5}+\underbrace{Cov(X_1,X_3)}_{=0.5}\}\\
&=\frac{2}{3}
\end{aligned}$$

と,

$$\begin{aligned}
V[Y]&=\frac{1}{9}V[X_1+X_2+X_3]\\
&=\frac{1}{9}\{V[X_1]+V[X_2]+V[X_3]+2Cov(X_1,X_2)+2Cov(X_2,X_3)+2Cov(X_3,X_1)\}\\
&=\frac{2}{3}
\end{aligned}$$

より, 相関係数は,

$$\frac{Cov(X_1,Y)}{\sqrt{V[X_1]}\sqrt{V[Y]}}=\frac{2/3}{\sqrt{2/3}}=\sqrt{\frac{2}{3}}\approx 0.8$$

となります。$\boxed{15}$の正解は⑤となります。

問 9（解答番号 $\boxed{16}$〜$\boxed{18}$ ）

[1]

ポアソン分布では, 母数 = 平均 = 分散 です。

$\boxed{16}$の正解は①となります。

[2]

イベントの参加者数を X と表しますと, X の平均と分散はともに 50 であり, また, 中心極限定理により, X は正規分布に従うと近似できます。よって,

$$P(X > 60) = P\left(\frac{X-50}{\sqrt{50}} > \frac{60-50}{\sqrt{50}}\right) \approx P(Z > 1.41) \approx 0.07$$

となります（Z は標準正規分布に従う確率変数)。$\boxed{17}$の正解は⑤となります。

[3]

事前登録をしない参加者数を Y, 参加者数全体の人数を W としますと, $W = Y + 30$ となります。[2] と同様に, Y は平均と分散が 20 の正規分布に従うと近似しますと, W は平均 50, 分散 20 の正規分布に従います。よって,

$$\begin{aligned}
& P(W > x) = 0.05 \\
\therefore \quad & P\left(\frac{W-50}{\sqrt{20}} > \frac{x-50}{\sqrt{20}}\right) = 0.05 \\
\therefore \quad & P\left(Z > \frac{x-50}{\sqrt{20}}\right) = 0.05 \quad (Z \text{ は標準正規分布に従う確率変数}) \\
\therefore \quad & \frac{x-50}{\sqrt{20}} = 1.645 \quad (\text{標準正規分布の上側 5 \%点}) \\
\therefore \quad & x \approx 57
\end{aligned}$$

となります。$\boxed{18}$の正解は③となります。

問 10（解答番号 $\boxed{19}$〜$\boxed{20}$ ）

[1]

$$\begin{aligned}
E[X^2] &= \int_{-\infty}^{\infty} x^2 \ f(x)dx \\
&= \int_{-1}^{0} x^2(x+1)dx + \int_{0}^{1} x^2(-x+1)dx \\
&= \left[\frac{1}{4}x^4 + \frac{1}{3}x^3\right]_{-1}^{0} + \left[-\frac{1}{4}x^4 + \frac{1}{3}x^3\right]_{0}^{1} \\
&= \frac{1}{6}
\end{aligned}$$

となります。$E[X] = 0$ より, $V[X] = E[X^2] - (E[X])^2 = \dfrac{1}{6}$ となります。

$\boxed{19}$の正解は①となります。

[2] a を第 1 四分位数としますと, $\displaystyle\int_{-\infty}^{a} f(x)dx = \frac{1}{4}$ を解くことになります。明らかに a は

$-1 \leq x < 0$ の範囲にありますので,

$$\int_{-\infty}^{a} f(x)dx = \int_{-1}^{a} (x+1)dx = \left[\frac{1}{2}x^2 + x\right]_{-1}^{a} = \frac{1}{2}a^2 + a + \frac{1}{2}$$

より, $\frac{1}{2}a^2 + a + \frac{1}{2} = \frac{1}{4}$ つまり $2a^2 + 4a + 1 = 0$ を解いて, $-1 \leq a < 0$ より, $a = \frac{-2+\sqrt{2}}{2} = -1 + \frac{\sqrt{2}}{2}$ となります。

$\boxed{20}$ の正解は③となります。

問 11（解答番号 $\boxed{21}$ ～ $\boxed{22}$ ）

[1]

分散に関する式 $V[X_i] = E[X_i^2] - (E[X_i])^2$ と, $E[X_i] = \mu$, $V[X_i] = \sigma^2$ から,

$$E[X_i^2] = (E[X_i])^2 + V[X_i] = \mu^2 + \sigma^2$$

となります。$\boxed{21}$ の正解は③となります。

[2]

（ア）には, 不偏分散が入ります。①か②になります。

（イ）に関して, $V[\bar{X}] = E[\bar{X}^2] - (E[\bar{X}])^2$ と, $E[\bar{X}] = \mu$, $V[\bar{X}] = \sigma^2/n$ から,

$$E[\bar{X}^2] = (E[\bar{X}])^2 + V[\bar{X}] = \mu^2 + \frac{\sigma^2}{n}$$

ですので,

$$E\left[\bar{X}^2 - \frac{\hat{\sigma}^2}{n}\right] = E[\bar{X}^2] - \frac{1}{n}E[\hat{\sigma}^2] = \mu^2 + \frac{\sigma^2}{n} - \frac{\sigma^2}{n} = \mu^2$$

となります。$\boxed{22}$ の正解は②となります。

問 13（解答番号 $\boxed{25}$ ～ $\boxed{27}$ ）

[1]

観測度数を表にまとめると, 以下になります。

	60～100	100～120	120～140	140～160	160～200	合計
夏季	18	42	24	11	1	96
冬季	13	25	23	19	16	96
合計	31	67	47	30	17	192

となります。求める期待度数は, $67 \times 96/192 = 33.5$ となります。

$\boxed{25}$ の正解は④となります。

[2]

上記の分割表を用いた, 独立性の χ^2 検定です。自由度は $(2-1) \times (5-1) = 4$ となります。

$\boxed{26}$ の正解は②となります。

[3]

自由度 4 の χ^2 分布の上側 5 %点 9.49 と比較して, 20.51 の方が大きいため, 帰無仮説「2 つの分布は同等である」を棄却します。

$\boxed{27}$ の正解は①となります。

【第10回：2017年6月実施】

問6（解答番号 $\boxed{13}$ ）

$$X = a + b + \varepsilon_1, \quad Y = a - b + \varepsilon_2$$

より，

$$\frac{X+Y}{2} = a + \frac{\varepsilon_1 + \varepsilon_2}{2}, \quad \frac{X-Y}{2} = b + \frac{\varepsilon_1 - \varepsilon_2}{2}$$

ですので，

$$\begin{aligned} V\left[\frac{X-Y}{2}\right] = V\left[b + \frac{\varepsilon_1 - \varepsilon_2}{2}\right] &= V\left[\frac{\varepsilon_1 - \varepsilon_2}{2}\right] \\ &= \frac{1}{4}(V[\varepsilon_1] + V[\varepsilon_2]) \qquad (\varepsilon_1\text{と}\varepsilon_2\text{は独立なので}) \\ &= \frac{1}{4}(\sigma^2 + \sigma^2) = \frac{\sigma^2}{2} \end{aligned}$$

となります。$\boxed{13}$の正解は③となります。

問9（解答番号 $\boxed{19}$〜$\boxed{20}$ ）
[1]

$$\begin{aligned} E[U] &= E[X+Y] = E[X] + E[Y] = 0 + 0 = 0 \\ E[V] &= E[X-Y] = E[X] - E[Y] = 0 - 0 = 0 \end{aligned}$$

です。また，X と Y は独立なので，

$$\begin{aligned} V[U] &= V[X+Y] = V[X] + V[Y] = \sigma_1^2 + \sigma_2^2 \\ V[Y] &= V[X-Y] = V[X] + V[Y] = \sigma_1^2 + \sigma_2^2 \end{aligned}$$

です。また，

$$\begin{aligned} Cov(U,V) = Cov(X+Y, X-Y) &= Cov(X+Y, X) - Cov(X+Y, Y) \\ &= Cov(X,X) + Cov(Y,X) - Cov(X,Y) - Cov(Y,Y) \\ &= V[X] - V[Y] = \sigma_1^2 - \sigma_2^2 \end{aligned}$$

です。よって，相関係数は，

$$r = \frac{Cov(U,V)}{\sqrt{V[U]}\sqrt{V[V]}} = \frac{\sigma_1^2 - \sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}\sqrt{\sigma_1^2 + \sigma_2^2}} = \frac{\sigma_1^2 - \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

となります。$\boxed{19}$の正解は④となります。

[2]

I : 正しいです。どちらも 0 です。

II : [1] の結果より, $\sigma_1^2 = \sigma_2^2$ のとき, $r = 0$ となります。U, V はともに正規分布に従いますが, その場合は, $r = 0$ ならば U と V は独立です（無相関ならば独立, というのは, いつも成り立つとは限りませんので, ご注意下さい)。正しいです。

III : ともに, 期待値 0, 分散 $\sigma_1^2 + \sigma_2^2$ の正規分布に従います。正しいです。

$\boxed{20}$の正解は⑤となります。

問 10（解答番号 $\boxed{21}$〜$\boxed{22}$ ）

[1]

W_1 は自由度 1 の χ^2 分布に従います。その上側 5 %点は 3.84 となります。

$\boxed{21}$の正解は④となります。

[2]

W_n は自由度 n の χ^2 分布に従います。

$n = 3$ のとき, 自由度 3 の χ^2 分布の上側 5 %点 7.81 より, $W_3 \geq 6$ となる確率は, 5 %より大きいです。

$n = 8$ のとき, 自由度 8 の χ^2 分布の上側 5 %点 15.51 より, $W_8 \geq 16$ となる確率は, 5 %より小さいです。

$\boxed{22}$の正解は②となります。

問 11（解答番号 $\boxed{23}$ ）

母平均を μ, 標本平均（母平均の推定値）を $\overline{X}$, 母分散を σ^2, 標本数を n とします。

まず, 母変動係数が 0.4 なので, $\dfrac{\sigma}{\mu} = 0.4$, つまり $\mu = \sigma/0.4$ です。また, 問題文の内容を式に表しますと,

$$Pr\left(-0.05 \leq \frac{\overline{X} - \mu}{\mu} \leq 0.05\right) = 0.95$$

となります。この分母の μ を $\sigma/0.4$ に置き換え, 整理しますと,

$$Pr\left(-\frac{1}{8} \leq \frac{\overline{X} - \mu}{\sigma} \leq \frac{1}{8}\right) = 0.95$$

となります。ところで $\overline{X}$ は期待値 μ, 分散 σ^2/n の正規分布に従いますので, 括弧内の不等式の辺々

に $\sqrt{n}$ を掛けて,

$$Pr\left(-\frac{\sqrt{n}}{8} \leq \underbrace{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}}_{N(0,1)\text{ に従う}} \leq \frac{\sqrt{n}}{8}\right) = 0.95$$

としますと, $\dfrac{\sqrt{n}}{8}$ は標準正規分布の上側 2.5 %点 1.96 に等しいことがわかります。よって $n \approx 250$ となります。

$\boxed{23}$ の正解は④となります。

（以上です）