# 第 7 問解答（統計学）

## $\boxed{\textbf{問題 1}}$

**1.** (1) (a) $\boxed{\text{(ア) 減少}}$, $\boxed{\text{(イ) 下}}$

(b) $\boxed{\text{(ウ) 0.5}}$

(c) $\boxed{\text{(エ) 完全平等線}}$, $\boxed{\text{(オ) 2}}$

(2) $x_1, x_2, \ldots, x_n$ は正の値をとる大きさ $n$ のデータで，それらを小さい方から大きさの順に並べたものを $x_{(1)} \leqq x_{(2)} \leqq \cdots \leqq x_{(n)}$ とし，この $k$ 番目までの和を $s_k = \sum_{i=1}^{k} x_{(i)}$ $(k = 1, 2, \ldots, n)$ とする．いま $xy$ 平面において，原点 O を $\mathrm{P}_0$，座標 $\left(\frac{k}{n}, \frac{s_k}{s_n}\right)$ に対応する点を $\mathrm{P}_k$ とし，これらの点 $\mathrm{P}_0, \mathrm{P}_1, \mathrm{P}_2, \ldots, \mathrm{P}_n$ を順に線分で結んだ折れ線 $L$ を考える．この折れ線を “ローレンツ曲線” とよぶ．

大きさ $n$ のデータ $x_1, x_2, \ldots, x_n$ のすべての組み合わせ $(x_i, x_j)$ $(i, j = 1, 2, \ldots, n)$ を考え，それらの差の絶対値を平均した $\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|$ と 標本平均 $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ の 2 倍との比

$$\frac{1}{2n^2\bar{x}}\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j| = \frac{1}{n^2\bar{x}}\sum_{i<j}|x_i - x_j|$$

を” ジニ係数” とよぶ.

(a) ローレンツ曲線： $(0,0)$, $(0.2, 0.05)$, $(0.4, 0.15)$, $(0.6, 0.30)$, $(0.8, 0.55)$, $(1,1)$ を結ぶ折れ線（図は省略）

$$\text{ジニ係数} = \frac{1+2+4+8+1+3+7+2+6+4}{5^2 \cdot 4} = \boxed{0.38}$$

(b) ローレンツ曲線： $(0,0)$, $(0.2, 0.10)$, $(0.4, 0.20)$, $(0.6, 0.30)$, $(0.8, 0.65)$, $(1,1)$ を結ぶ折れ線（図は省略）

$$\text{ジニ係数} = \frac{0+0+5+5+0+5+5+5+5+0}{5^2 \cdot 4} = \boxed{0.30}$$

**2.** (1) $\boxed{\text{(ア) 0.25}}$, $\boxed{\text{(イ) 0.35}}$, $\boxed{\text{(ウ) 0.20}}$, $\boxed{\text{(エ) 0.20}}$

$\boxed{\text{(オ) 階級}}$, $\boxed{\text{(カ) 階級値}}$, $\boxed{\text{(キ) 度数}}$, $\boxed{\text{(ク) 相対度数}}$

(2) 平均値 $= 8.5 \times 0.25 + 9.5 \times 0.35 + 10.5 \times 0.20 + 12.0 \times 0.20 = \boxed{9.95}$

(3) 階級 8 〜 9 の柱の高さ：0.25，階級 9 〜 10 の柱の高さ：0.35,

階級 10 〜 11 の柱の高さ：0.20，階級 11 〜 13 の柱の高さ：0.01

となるヒストグラム（図は省略）

# 第 7 問解答（統計学）

## $\boxed{\textbf{問題 2}}$

(1) 周辺確率 $p(x) = \sum_y p(x, y)$ の表

| $x$ | 3 | 6 |
|---|---|---|
| $p(x)$ | 0.51 | 0.49 |

周辺確率 $p(y) = \sum_x p(x, y)$ の表

| $y$ | $-2$ | 0 | 2 | 4 |
|---|---|---|---|---|
| $p(y)$ | 0.25 | 0.19 | 0.26 | 0.30 |

(2) $E(Y) = -2 \times 0.25 + 0 \times 0.19 + 2 \times 0.26 + 4 \times 0.30 = \boxed{1.22}$

(3) 条件付確率 $p(y|x = 3) = \dfrac{p(3, y)}{p(3)}$ の表

| $y$ | $-2$ | 0 | 2 | 4 |
|---|---|---|---|---|
| $p(y\|x = 3)$ | 0.490 | 0.196 | 0.314 | 0.000 |

$E(Y|X = 3) = -2 \times 0.490 + 0 \times 0.196 + 2 \times 0.314 + 4 \times 0.000 = \boxed{-0.352}$

$E(Y^2|X = 3) = (-2)^2 \times 0.490 + 0^2 \times 0.196 + 2^2 \times 0.314 + 4^2 \times 0.000 = 3.216$

$\therefore \mathrm{var}(Y|X = 3) = E(Y^2|X = 3) - \{E(Y|X = 3)\}^2 = 3.216 - (-0.352)^2 = \boxed{3.092}$

## 第 7 問解答（統計学）

**問題 3**

**1.** (1) $Y_s = X_1 + X_2 + \cdots + X_s$, $Y_k - Y_s = X_{s+1} + X_{s+2} + \cdots + X_k$ なので，両者は $\boxed{\text{独立である}}$

(2) $\Pr(Y_2 = 2) = \Pr(X_1 = 1, X_2 = 1) = \Pr(X_1 = 1)\Pr(X_2 = 1) = \boxed{\dfrac{1}{4}}$

(3) $\Pr(Y_5 = 3) = \Pr(5\text{ 回の試行で表が }4\text{ 回}) = \dbinom{5}{4}\left(\dfrac{1}{2}\right)^4\left(\dfrac{1}{2}\right) = \boxed{\dfrac{5}{32}}$

(4) $\Pr(Y_2 = 2, Y_5 = 3) = \Pr(Y_2 = 2, Y_5 - Y_2 = 1) = \Pr(Y_2 = 2)\Pr(X_3 + X_4 + X_5 = 1) = \dfrac{1}{4} \times \dbinom{3}{2}\left(\dfrac{1}{2}\right)^2\left(\dfrac{1}{2}\right) = \boxed{\dfrac{3}{32}}$

**2.** (1) $\boxed{\text{(ア) 奇}}$

(2) $\boxed{\text{(イ) } 1 - F(x)}$

(3) $\boxed{\text{(ウ) } k}$ , $\boxed{\text{(エ) カイ 2 乗分布}}$ , $\boxed{\text{(オ) } k}$

# 第 8 問解答（統計学）

**問題 1**

(1) $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 167.7 - 20.45 \times 6,7 = \boxed{30.69}$

$$\hat{\beta} = \frac{\sum_j x_j y_j - 10\bar{x}\bar{y}}{\sum_j x_j^2 - 10\bar{x}^2} = \frac{11647 - 10 \times 6.7 \times 167.7}{469 - 10 \times 6.7^2} = \boxed{20.45}$$

(2)

$\hat{\beta}$ の標準誤差を $s(\hat{\beta})$ と表すと

$$t = \frac{\hat{\beta} - \beta}{s(\hat{\beta})} \sim t(8)$$

であるので，$H_0 : \beta = 0$ を $H_1 : \beta \neq 0$ を有意水準 5% で検定するには，

$$|t| = \left|\frac{\hat{\beta}}{s(\hat{\beta})}\right| > t_{0.025}(8) = 2.306$$

のとき，$H_0$ を棄却すればよい．
実際，$|t| = \dfrac{20.45}{2.87} = 7.125 > 2.306$ なので，$H_0$ は棄却される．

(3) 予測値 $= \hat{\alpha} + \hat{\beta} \times 12 = 30.69 + 20.45 \times 12 = \boxed{279.09}$

(4)

はじめの 5 社の広告費を $x_{1i}$，残りの 5 社の広告費を $x_{2i}$ で表し，

$$x_{1i} = \begin{cases} x_i & (i = 1, 2, \ldots, 5) \\ 0 & (i = 6, 7, \ldots, 10) \end{cases} \qquad x_{2i} = \begin{cases} 0 & (i = 1, 2, \ldots, 5) \\ x_i & (i = 6, 7, \ldots, 10) \end{cases}$$

と定義する．これらを使い，回帰モデル

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i, \quad i = 1, 2, \ldots, 10$$

を設定すればよい．ここで，$\beta_1$ ははじめの 5 社の広告効果，$\beta_2$ は残りの 5 社の広告効果を表す母数である．

（別解）

はじめの 5 社，残りの 5 社のどちらのグループに属するかを表すダミー変数

$$\delta_i = \begin{cases} 1 & (i = 1, 2, \ldots, 5) \\ 0 & (i = 6, 7, \ldots, 10) \end{cases}$$

を導入し，回帰モデル

$$y_i = \alpha + \beta x_i + \gamma \delta_i x_i + \epsilon_i, \quad i = 1, 2, \ldots, 10$$

を設定すればよい．ここで，$\beta$ は 2 つのグループ共通の広告効果，$\gamma$ はグループの違いによる広告効果を表す母数である．

## 第 8 問解答（統計学）

### 問題 2

(1) $\lambda = 2.3$ であるので

$0$ 件の確率 $= e^{-2.3} = \boxed{0.1000}$

$1$ 件の確率 $= 2.3 \times e^{-2.3} = \boxed{0.2300}$

$2$ 件の確率 $= \dfrac{2.3^2}{2} \times e^{-2.3} = \boxed{0.2645}$

$3$ 件以上の確率 $= 1 - (0.1000 + 0.2300 + 0.2645) = \boxed{0.4055}$

(2) $0$ 件の期待日数 $= 60 \times 0.1000 = \boxed{6.00}$

$1$ 件の期待日数 $= 60 \times 0.2300 = \boxed{13.80}$

$2$ 件の期待日数 $= 60 \times 0.2645 = \boxed{15.87}$

$3$ 件以上の期待日数 $= 60 \times 0.4055 = \boxed{24.33}$

(3)

| 請求件数 | 0 | 1 | 2 | 3 以上 |
|---|---|---|---|---|
| 実測日数 | 7 | 16 | 21 | 16 |
| 期待日数 | 6.00 | 13.80 | 15.87 | 24.33 |

この結果からカイ 2 乗値 $\chi^2$ を計算し,

$$\chi^2 > \chi_{0.05}(3) = 7.81$$

なら，平均 2.3 のポアソン分布にしたがうという仮説は有意水準 5% で棄却される.

実際

$$\begin{aligned}\chi^2 &= \frac{(7-6.00)^2}{6.00} + \frac{(16-13.80)^2}{13.80} + \frac{(21-15.87)^2}{15.87} + \frac{(16-24.33)^2}{24.33} \\ &= 5.03 \not> 7.82\end{aligned}$$

よって，平均 2.3 のポアソン分布にしたがうという仮説は棄却されない.

## 第 8 問解答（統計学）

**問題 3**

(1) $\boxed{\text{(ア) 自己共分散}}$ , $\boxed{\text{(イ) 自己相関係数}}$ , $\boxed{\text{(ウ) } c_h}$ , $\boxed{\text{(エ) } c_0}$

(2) $\boxed{\text{(オ) ホワイトノイズ}}$ , $\boxed{\text{(カ) 自己回帰}}$ , $\boxed{\text{(キ) } < 1}$

(3) $\boxed{\text{(ク) } \dfrac{b}{1-a}}$ , $\boxed{\text{(ケ) } \dfrac{\sigma^2}{1-a^2}}$ , $\boxed{\text{(コ) } \dfrac{a^h\sigma^2}{1-a^2}}$ , $\boxed{\text{(サ) } a^h}$