今野　(2008 年 1 月 25 日)

確率統計と情報処理・演習の試験問題（試験時間は 75 分）

解答上の注意

特別な指示がある場合を除き，答えが合っているかどうかよりも解答の途中過程の論理的展開を重視して採点をします．以下の点に留意して解答を作成すること．

(1) 解答の途中過程は丁寧に記述すること．

(2) 講義で述べたことは設問中で証明することを求めていない場合には証明なしに用いてよい．しかし，なにをどのように用いたかを可能な限り明示すること．

(3) 各自が理解していることを採点者にわかるように解答を作成することを心がけること．

(4) 等号の使い方に注意すること．

(5) どの問題を解答しているかが明示されていれば，解答は問題番号順でなくともよい．

答案は来年度開講の「統計解析」の第 1 回目の講義で返却をする．受講しないものは研究室まで取りに来てください．

問題 1　$n \geq 2$ とする．n 個のデータの値を $x_1, x_2, \ldots, x_n$ とする．このとき，以下の問いに答えよ．

(1) 実数 $a \neq 0$ と b に対し，$y_i = ax_i + b\,(i = 1, 2, \ldots, n)$ とし，

$$\bar{y}_n = \frac{1}{n}\sum_{i=1}^{n} y_i, \qquad \bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$$

と定める．このとき，関係式

$$\bar{y}_n = a\bar{x}_n + b$$

を示せ．

(2) 実数 c に対して，

$$g(c) = \sum_{i=1}^{n}(x_i - c)^2,$$

と定めたとき，不等式

$$g(c) \geq \sum_{i=1}^{n}(x_i - \bar{x}_n)^2$$

を示せ．等号が成立する場合には，その条件を求めよ．

(3) $\sum_{i=1}^{n}(x_i - \bar{x}_n)^2$ と $\sum_{i=1}^{n}(y_i - \bar{y}_n)^2$ の間に成立する関係式を求め，それを示せ．

解答と配点　(1) 10 (2) 10 (3) 10 計 30

(1) $\sum_{i=1}^{n} x_i = n\bar{x}_n$ に注意して，

$$\bar{y}_n = \sum_{i=1}^{n} y_i \quad = \quad \frac{1}{n}\sum_{i=1}^{n}(ax_i + b) = \frac{1}{n}\left\{a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} 1\right\} = \frac{1}{n}(an\bar{x}_n + nb) = a\bar{x}_n + b$$

(2) $\bar{x}_n$ を差し繰し，$\sum_{i=1}^{n}(x_i - \bar{x}_n) = 0$ であることに注意すれば，

$$\begin{aligned} g(c) &= \sum_{i=1}^{n}(x_i - \bar{x}_n + \bar{x}_n - c)^2 = \sum_{i=1}^{n}\{(x_i - \bar{x}_n)^2 + (\bar{x}_n - c)^2 + 2(x_i - \bar{x}_n)(\bar{x}_n - c)\} \\ &= \sum_{i=1}^{n}(x_i - \bar{x}_n)^2 + \sum_{i=1}^{n}(\bar{x}_n - c)^2 + 2(\bar{x}_n - c)\sum_{i=1}^{n}(x_i - \bar{x}_n) = \sum_{i=1}^{n}(x_i - \bar{x}_n)^2 + n(\bar{x}_n - c)^2 \end{aligned}$$

しかし，$n(\bar{x}_n - c)^2 \geq 0$ なので，

$$g(c) \geq \sum_{i=1}^{n}(x_i - \bar{x}_n)^2$$

等号成立は

$$n(\bar{x}_n - c)^2 = 0 \Leftrightarrow c = \bar{x}_n$$

(3) $\bar{y}_n = a\bar{x}_n + b$ に注意して，

$$\sum_{i=1}^{n}(y_i - \bar{y}_n)^2 \quad = \quad \sum_{i=1}^{n}(ax_i + b - (a\bar{x}_n + b))^2 = a^2\sum_{i=1}^{n}(x_i - \bar{x}_n)^2$$

$\boxed{\text{問題 2}}$ 確率変数 X は閉区間 $[0, 1]$ 上の一様分布にしたがっているとする．すなわち，任意の実数 a, b に対し，

$$\mathbb{P}(a < X \leq b) = \int_a^b f_X(x)\,dx, \qquad f_X(x) = \begin{cases} 1 & (0 \leq x \leq 1), \\ 0 & (\text{その他}) \end{cases}$$

となる．このとき，以下の問いに答えよ．ただし，積分の計算においては，$\boxed{\text{問題 3}}$ のヒントを証明なしで用いてよい．

(1) X の期待値

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx,$$

を計算せよ．

(2) 実数上で定義された関数である X の分布関数

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$$

を求め，$F_X(x)$ のグラフを作図せよ．

$\boxed{\text{解答と配点}}$ (1) 10 (2) 10 計 20

(1)

$$\begin{aligned}\mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_{-\infty}^{0} x f_X(x)\,dx + \int_{0}^{1} x f_X(x)\,dx + \int_{1}^{\infty} x f_X(x)\,dx \\ &= \int_0^1 x f_X(x)\,dx = \int_0^1 x\,dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2}\end{aligned}$$

(2) $x < 0$ のとき，$f_X(x) = 0$ となるので，

$$F_X(x) = 0$$

$0 \leq x < 1$ のとき，

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \int_{-\infty}^{0} f_X(t)\,dt + \int_0^x f_X(t)\,dt = [t]_0^x = x$$

$x \geq 1$ のとき，

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \int_{-\infty}^{0} f_X(t)\,dt + \int_0^1 f_X(t)\,dt + \int_1^x f_X(t)\,dt = [t]_0^1 = 1$$

したがって，

$$F_X(x) = \begin{cases} 0 & (x < 0), \\ x & (0 \leq x < 1), \\ 1 & (x \geq 1) \end{cases}$$

$\boxed{\text{問題 3}}$ 確率変数 Z は標準正規分布に従うとする．すなわち，任意の実数 $a, b\,(a < b)$ に対して，

$$\mathbb{P}(a < Z \leq b) = \int_a^b f_Z(z)\,dz, \qquad f_Z(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}, \qquad \int_{-\infty}^{\infty} f_Z(z)\,dz = 1$$

である．また，$\mathbb{P}(Z \le b) = \lim_{a\to-\infty} \mathbb{P}(a < Z \le b)$ と $\mathbb{P}(Z > a) = \lim_{b\to\infty} \mathbb{P}(a < Z \le b)$ とする．

(1) 任意の正の実数 c に対して，

$$\mathbb{P}(Z \le -c) = \mathbb{P}(Z > c)$$

を示せ．

(2) $0 < \alpha < 1/2$ なる数 α に対して，実数 k_α は方程式

$$\int_{-\infty}^{k_\alpha} f_Z(z)\,dz = 1 - \alpha$$

をみたす点とする．このとき，

$$k_\alpha > 0$$

を示せ．

(3) $0 < \alpha < 1/2$ なる数 α に対して，実数 k_α は方程式

$$\int_{-\infty}^{k_\alpha} f_Z(z)\,dz = 1 - \alpha$$

をみたす点とする．このとき，確率

$$\mathbb{P}(0 < Z \le k_\alpha)$$

を求めよ．

ヒント：以下の積分の性質や公式は証明なしに用いてよい．

- 実数 a, b, c に対して，$\int_a^c f_Z(z)\,dz = \int_a^b f_Z(z)\,dz + \int_b^c f_Z(z)\,dz$
- $\int_a^b f_Z(z)\,dz = -\int_b^a f_Z(z)\,dz$
- $a < b$ とする．$g(x) > 0\,(a \le x \le b)$ のとき，$\int_a^b g(x)\,dx > 0$
- 積分の変数変換の公式

$\boxed{\text{解答と配点}}$ (1) 10 (2) 10 (3) 10 計 30

(1) $w = -z$ とおけば，$f_Z(-z) = f_Z(w)$ に注意して，

$$\begin{aligned}
\mathbb{P}(Z \le -c) &= \lim_{a\to-\infty} \mathbb{P}(a < Z \le -c) = \lim_{a\to-\infty} \int_a^{-c} f_Z(z)\,dz = \lim_{a\to-\infty} \int_{-a}^{c} f_Z(-w)\,(-dw) \\
&= \lim_{a\to-\infty} \int_{-a}^{c} -f_Z(w)\,dw = \lim_{a\to-\infty} \int_{c}^{-a} f_Z(w)\,dw = \lim_{b\to\infty} \int_c^b f_Z(w)\,dw = \lim_{b\to\infty} \mathbb{P}(c < Z \le b) \\
&= \mathbb{P}(c < Z)
\end{aligned}$$

(2) $c = 0$ とおけば，(1) より

$$\frac{1}{2} = \mathbb{P}(Z \le 0) = \int_{-\infty}^{0} f_Z(z)\,dz$$

となる．さらに，$0 < \alpha < 1/2$ なので，

$$\int_{-\infty}^{k_\alpha} f_Z(z)\,dz > \frac{1}{2} = \int_{-\infty}^{0} f_Z(z)\,dz$$

したがって，

$$0 < \int_{-\infty}^{k_\alpha} f_Z(z)\,dz - \int_{-\infty}^{0} f_Z(z)\,dz = \int_0^{k_\alpha} f_Z(z)\,dz$$

となる．$f_Z(z) > 0$ より，$k_\alpha > 0$ となる．

(3)

$$\mathbb{P}(0 < Z < k_\alpha) = \int_0^{k_\alpha} f_Z(z)\,dz = \int_{-\infty}^{k_\alpha} f_Z(z)\,dz - \int_{-\infty}^{0} f_Z(z)\,dz = 1 - \alpha - \frac{1}{2} = \frac{1}{2} - \alpha$$

$\boxed{\text{問題 4}}$ 　標本空間を Ω とし，Ω の部分集合（事象）を入力とする実数値関数 $\mathbb{P}(\cdot)$ で，つぎの条件をみたすものを考える[1]：

(P1) すべての事象 A に対して，$0 \le \mathbb{P}(A) \le 1$

(P2) $\mathbb{P}(\Omega) = 1$

(P3) 事象 A と B が互いに排反[2]のとき，$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ が成立

このとき，以下の問いに答えよ．ただし，A と B は事象とし，A^c は A の補事象である．

(1) 事象 A に対し，$\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ を示せ．

(2) $\mathbb{P}(A \cup B) \le \mathbb{P}(A) + \mathbb{P}(B)$ を示せ．

$\boxed{\text{解答と配点}}$ 　(1) 10 (2) 10 計 20

(1) 補事象の定義から $A \cap A^c = \emptyset$ となるので，A と A^c は互いに排反．さらに，$A \cup A^c = \Omega$ より

$$\begin{aligned} 1 &= \mathbb{P}(\Omega) && \text{（確率の性質 (P2) より）} \\ &= \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c) && \text{（確率の性質 (P3) より）} \end{aligned}$$

(2) $A \cap B^c$, $A \cap B$, $A^c \cap B$ は互いに排反で，$(A \cap B^c) \cup (A \cap B) \cup (A^c \cap B) = A \cup B$ に注意して，P3 を用いれば，

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

しかし，確率の性質 (P1) より $\mathbb{P}(A \cap B) \ge 0$ なので，

$$\mathbb{P}(A \cup B) \le \mathbb{P}(A) + \mathbb{P}(B)$$

となる．

$\boxed{\text{解答と配点}}$

[1] $\mathbb{P}$ を教科書では Pr と書いている．
[2] すなわち，$A \cap B = \emptyset$ が成立する．

成績について

得点	0 ∼ 9	10 ∼ 29	30 ∼ 49	50 ∼ 69	70 ∼ 100
成績	D	C	B	A	A^+

得点分布　平均点 = 35.4，中央値 = 40.0、標準偏差 = 12.8

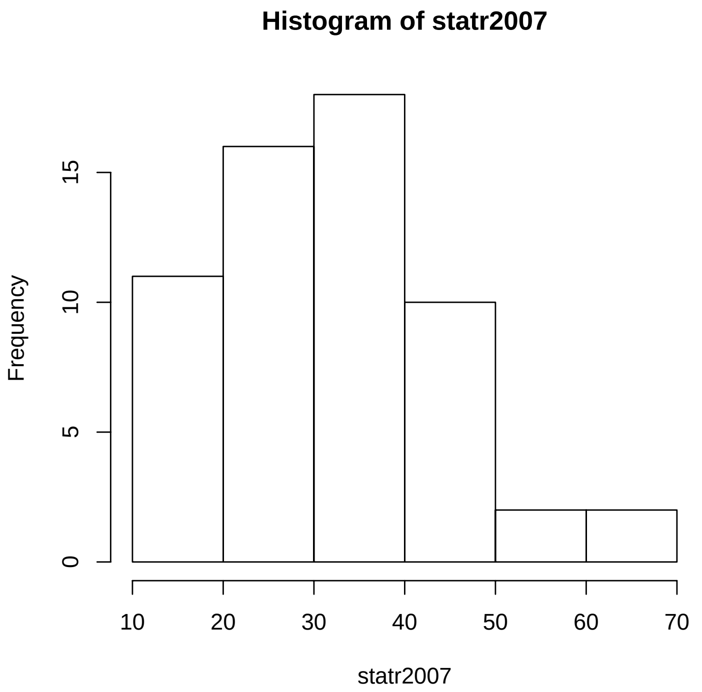

Figure 1: This is a figure