今野　(2009 年 1 月 20 日)

情報統計学の試験問題（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) 設問中で証明することを求めている場合をのぞき，講義で述べたこと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．

問題 1　確率変数 X は確率密度関数

$$f_X(x|\theta) = \begin{cases} \dfrac{1}{\theta} & (0 \le x \le \theta) \\ 0 & (\text{その他}) \end{cases}$$

からの標本の大きさ 1 のランダム標本とする．ただし，$\theta > 0$ とする．
(1) X の分布関数 $F_X(x|\theta) = \int_{-\infty}^{x} f_X(t|\theta)\,dt$ を求めよ．
(2) $X = x\,(x > 0)$ が観測されたときの尤度関数 $L(\theta|x)$ のグラフを書け．ただし，横軸は θ で，縦軸は $L(\theta|x)$ とする．グラフから，最尤推定値 x を求めよ（証明の必要はない）．
(3) θ の推定量 $2X$ は θ の不偏推定量かどうかを調べよ．
(4) θ の推定量 $2X$ の平均二乗誤差 $MSE_{2X}(\theta) = \mathbb{E}[(2X-\theta)^2]$ を求め，そのグラフを描け．ただし，横軸は θ で，縦軸は $MSE_{2X}(\theta)$ とする．
(5) $X_1, X_2, \ldots, X_n\,(n \ge 2)$ を確率密度関数 $f_X(x|\theta)$ からのランダム標本とし，$X_{(n)}$ を $X_1, \ldots, X_n$ を小さい順に並び替えたときの最大の確率変数とする．このとき，$X_{(n)}$ の確率密度関数 $f_{X_{(n)}}(x|\theta)$ を求め，$\int_{-\infty}^{\infty} f_{X_{(n)}}(x|\theta)\,dx = 1$ を確認せよ．

ヒント　(a) $X_1, X_2, \ldots, X_n$ を分布関数 $F_X(x)$ と確率密度関数 $f_X(x)$ をもつ連続型分布からの標本の大きさが n のランダム標本とする．$X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ を標本の順序統計量としたとき，$X_{(j)}, j = 1, 2, \ldots, n,$ の確率密度関数は

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x)[F_X(x)]^{j-1}[1-F_X(x)]^{n-j},$$

となる．

解答と配点　(1) 10 (2) 10 (3) 10 (4) 15 (5) 15 計 60
(1) $x < 0$ のとき，$f_X(x|\theta) = 0$ より

$$F_X(x|\theta) = \int_{-\infty}^{x} f_X(t|\theta)\,dt = 0.$$

$0 \le x \le \theta$ のとき，

$$F_X(x|\theta) = \int_{-\infty}^{x} f_X(t|\theta)\,dt = \int_{-\infty}^{0} f_X(t|\theta)\,dt + \int_{0}^{x} f_X(t|\theta)\,dt = \frac{x}{\theta}.$$

$x > \theta$ のとき，

$$F_X(x|\theta) = \int_{-\infty}^{x} f_X(t|\theta)\,dt = \int_{-\infty}^{0} f_X(t|\theta)\,dt + \int_{0}^{\theta} f_X(t|\theta)\,dt + \int_{\theta}^{x} f_X(t|\theta)\,dt = 1.$$

したがって，

$$F_X(x|\theta) = \begin{cases} 0 & x < 0; \\ \frac{x}{\theta} & 0 \le x \le \theta; \\ 1 & x > 0. \end{cases}$$

(2) 尤度関数 $L(\theta|\, x)$ は，$x\,(x>0)$ を固定したとき，θ を変数としてみることになる．したがって，

$$L(\theta|\, x) = \begin{cases} \frac{1}{\theta} & \theta \ge x; \\ 0 & \theta < x. \end{cases}$$

となるので，$\theta = x$ で $L(\theta|\, x)$ は最大となる．よって，最尤推定値は x である．
(3) θ の推定量 $2X$ は θ の不偏推定量であるためには，$\mathbb{E}[2X] = \theta\,(\forall\theta > 0$ を確認すればよい：

$$\mathbb{E}[2X] \;=\; \int_{-\infty}^{\infty} 2x f_X(x|\,\theta)\,dx = \int_0^{\theta} \frac{2x}{\theta}\,dx = \frac{2}{\theta}\int_0^{\theta} x\,dt = \frac{2}{\theta}\left[\frac{t^2}{2}\right]_0^{\theta} = \theta.$$

よって，$2X$ は θ の不偏推定量である．
(4) $2X$ は θ の不偏推定量より

$$\mathbb{E}[(2X-\theta)^2] = \mathbb{VAR}[2X] = 4\mathbb{VAR}[X] = 4\{\mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2$$

しかし，

$$\begin{aligned} \mathbb{E}[X] \;&=\; \frac{\theta}{2}; \\ \mathbb{E}[X^2] \;&=\; \int_{-\infty}^{\infty} x^2 f_X(x)\,dx = \int_0^{\infty} \frac{x^2}{\theta}\,dx = \theta^2 \int_0^1 t^2\,dt = \theta^2[\frac{t^3}{3}]_0^1 = \frac{\theta^2}{3}. \end{aligned}$$

よって，

$$MSE_{2X}(\theta) = 4\left\{\frac{\theta^2}{3} - \left(\frac{\theta}{2}\right)^2\right\} = \frac{\theta^2}{3}.$$

(5) ヒント (a) と (1) より $0 < x < \theta$ のとき，

$$\begin{aligned} f_{X_{(n)}}(x|\,\theta) \;&=\; \frac{n!}{(n-1)!} f_X(x|\,\theta)[F_X(x|\,\theta)]^{n-1} \\ &=\; n\frac{1}{\theta}\left(\frac{x}{\theta}\right)^{n-1} = \frac{nx^{n-1}}{\theta^n}. \end{aligned}$$

$x \notin (0, \theta)$ のときは，$f_X(x|\,\theta) = 0$ なので，$f_{X_{(n)}}(x|\,\theta) = 0$ は明らか．したがって，

$$f_{X_{(n)}}(x|\,\theta) = \begin{cases} \dfrac{nx^{n-1}}{\theta^n}, & 0 < x < \theta; \\ 0, & \text{その他}. \end{cases}$$

$t = x/\theta$ とおけば，

$$\begin{aligned} \int_{-\infty}^{\infty} f_{X_{(n)}}(x|\,\theta)\,dx \;&=\; \int_0^{\theta} \frac{x^{n-1}}{\theta^n}\,dx \\ &=\; n\int_0^1 t^{n-1}\,dt = n\left[\frac{t^n}{n}\right]_0^1 = 1. \end{aligned}$$

問題 2　$\{X_n\}_{n=1}^{\infty}$ を確率変数列とし，各 X_n は確率関数

$$f_{X_n}(x) = \begin{cases} \frac{1}{n} & (x = c-n), \\ 1 - \frac{1}{n} & (x = c), \\ 0 & (\text{その他}) \end{cases}$$

を持つとする．ただし，c は定数とする．このとき，以下の問いに答えよ．
(1) 期待値の定義に従い，$\mathbb{E}[X_n]$ を求めよ．
(2) X_n の分布関数 $F_{X_n}(x) = \mathbb{P}(X_n \le x)\,(x \in \mathbb{R})$ を求め，そのグラフを描け．
(3) 与えられた正の数 $\epsilon > 0$ に対して，十分大きな正の整数 n をとり，$n > \epsilon$ とする．このとき，$\mathbb{P}(X_n > c+\epsilon)$ と $\mathbb{P}(X_n < c-\epsilon)$ の値を求めよ．
(4) $X_n \xrightarrow{P} c$ を示せ．
ヒント (b) $n \to \infty$ のとき，$\{X_n\}$ が定数 c に**確率収束**するとは，任意の正数 ϵ に対して

$$\lim_{n\to\infty} P(|X_n - c| > \epsilon) = 0$$

をみたすときをいう．

$\boxed{\text{解答と配点}}$　(1) 10 (2) 10 (3) 10 (4) 10 計 40

(1) 期待値の定義より

$$\mathbb{E}[X_n] = (c-n)f_{X_n}(c-n) + cf_{X_n}(c) = (c-n)\frac{1}{n} + c\left(1 - \frac{1}{n}\right) = c - 1.$$

(2) X_n の取りうる値は $c-n$ または c なので，$x < c-n$ のとき，

$$F_{X_n}(x) = 0$$

となる．$c - n \leq x < c$ のとき，

$$F_{X_n}(x) = f_{X_n}(c-n) = \frac{1}{n}.$$

$x \geq c$ のとき，

$$F_{X_n}(x) = f_{X_n}(c-n) + f_{X_n}(c) = \frac{1}{n} + 1 - \frac{1}{n} = 1.$$

よって，

$$F_{X_n}(x) = \begin{cases} 0 & x < c-n; \\ \dfrac{1}{n} & c-n \leq x < c; \\ 1 & x \geq c. \end{cases}$$

(3) $\dfrac{1}{n} = F_{X_n}(c-n) \leq \mathbb{P}(X_n < c-\epsilon) \leq \lim_{x\uparrow c} F_{X_n}(x) = \dfrac{1}{n}$ より，

$$\mathbb{P}(X_n < c - \epsilon) = \frac{1}{n}.$$

また，

$$\mathbb{P}(X_n > c+\epsilon) = 1 - \mathbb{P}(X_n \leq c+\epsilon) = 1 - F_{X_n}(c+\epsilon) = 0.$$

(4) 任意の正の数 $\epsilon > 0$ に対して，

$$\{|X_n - c| > \epsilon\} = \{X_n > c+\epsilon\} \cup \{X_n < c-\epsilon\} \quad \text{かつ} \quad \{X_n > c+\epsilon\} \cap \{X_n < c-\epsilon\} = \emptyset$$

なので，

$$\mathbb{P}(|X_n - c| > \epsilon) = \mathbb{P}(X_n > c+\epsilon) + \mathbb{P}(X_n < c-\epsilon).$$

ここで，$n > \epsilon$ とすれば，(3) より

$$\mathbb{P}(|X_n - c| > \epsilon) = \mathbb{P}(X_n > c+\epsilon) + \mathbb{P}(X_n < c-\epsilon) = \frac{1}{n}$$

よって，

$$\lim_{n\to\infty} \mathbb{P}(|X_n - c| > \epsilon) = \lim_{n\to\infty} \frac{1}{n} = 0$$

なので，X_n は c に確率収束する．